Identification and Analyses of Variants Associated with COVID-19 from Non-invasive Prenatal Testing in Slovak Population

Natalia Forgacova[1], Juraj Gazdarica[2,3,4], Jaroslav Budis[1,3,4], Martina Sekelska[5,6] and Tomas Szemes[1,2,3]

[1]Comenius University Science Park, Ilkovicova 8, Bratislava, 841 04, Slovakia
[2]Faculty of Natural Sciences, Comenius University, Bratislava, 841 04, Slovakia
[3]Geneton Ltd., Bratislava, 841 04, Slovakia
[4]Slovak Centre of Scientific and Technical Information, Bratislava, 811 04, Slovakia
[5]Trisomy Test Ltd., Bratislava, 841 04, Slovakia
[6]Medirex Inc., Bratislava, 821 06, Slovakia

Abstract. Since December 2019, coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has rapidly spread throughout the world and caused a large global pandemic which drastically changed our everyday lives. As the COVID-19 pandemic progressed, a number of its characteristics showed enormous inter-individual and inter-population differences. Earlier genome-wide association studies (GWAS) have identified potential key genes and genetic variants associated with the risk and prognosis of COVID-19, but the underlying biological interpretation is largely unclear. Our previous work described genomic data generated through non-invasive prenatal testing (NIPT) based on low-coverage massively parallel whole-genome sequencing of total plasma DNA of pregnant women in Slovakia as a valuable source of population specific data. In the present study, we have performed a literature search of studies and used NIPT data to determine the population allele frequency of risk COVID-19 variants that have been reported in GWAS studies to date. We also focused on variants located in the *ACE2* gene, encoding angiotensin-converting enzyme 2 (*ACE2*), which is hypothesized to be a possible genetic risk factor for SARS-CoV-2 infection. Allele frequencies of identified variants were compared with six world populations from the gnomAD database to detect significant differences between populations. We interpreted variants and searched for functional consequences and clinical significance of variants using publicly available databases. Finally, 2 COVID-19 risk variants were found that showed statistically significant differences in population allele frequencies - rs383510 and rs1801274.

1 Introduction

The Coronavirus Disease (COVID-19), caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), is a complex, highly infectious disease involving the respiratory, immune, cardiovascular, gastrointestinal, and neurological systems [1–4]. The first case was registered in Wuhan, Hubei Province of China in December 2019, and it has rapidly evolved into a global pandemic [5]. At the time of the writing (June 2021), there have been more than 177 million confirmed cases and 3.8 million deaths worldwide (in Slovakia more than 390 000 people were infected, with the total deaths exceeding 12 000) (https://origin-coronavirus.jhu.edu/map.html).

Although the mortality rate of COVID-19 (ranges between 1-7%) is lower than that of the other two types of coronaviruses, severe acute respiratory syndrome (SARS-CoV) and the Middle East respiratory syndrome (MERS-CoV), the rate of human-to-human transmission is higher, as respiratory droplets and close contact can primarily transmit it [4,6–10]. COVID-19 presents a broad spectrum of varied clinical manifestations, from asymptomatic or mild symptoms to serious health outcomes leading to death [11,12]. Even though the symptoms are highly heterogeneous, the most commonly observed in the large majority of infected persons are fever, cough, severe headache, muscle pain, fatigue, myalgia, shortness of breath, chest tightness, and loss of taste or smell [13–18]. Besides, several minor symptoms such as gastrointestinal complications, including nausea, vomiting, and diarrhea, have also been reported [19]. In severe cases, breathing difficulties with dyspnea occur, with acute respiratory distress syndrome (ARDS) being the most serious complication [20]. It is likely that a mixture of genetic and nongenetic factors interplays between virus and host genetic background and determines the severity of COVID-19 outcome. Advanced age, male sex, some ethnicity and blood type such as A and AB0 blood types, smoking, hypertension, diabetes mellitus, obesity, cardiovascular, respiratory, and kidney disease or cancer have been identified as risk factors associated with a higher risk of death COVID-19 [12,21–27]. Nevertheless, these factors do not explain the main pathogenesis of COVID-19. Therefore, the host's genetic variations may partly provide novel insights into pathological mechanisms underlying COVID-19. Recently, genome-wide association studies (GWAS) have been performed to uncover genetic risk factors associated with the diagnosis and prognosis of COVID-19; however, the biological interpretation of their findings has not yet been fully clarified [28–32].

Non-invasive prenatal testing (NIPT) based on low-coverage massively parallel whole-genome sequencing of plasma DNA from pregnant women generates a large amount of data that provides the resources to investigate human genetic variations in the population. In our previous studies, we described the re-use of the data from NIPT for genome-scale population specific frequency determination of small DNA variants [33] and CNVs [34]. Since pregnant women represent a relatively standard sample of the local female population, we assumed this NIPT data could also be used in the population study of COVID-19.

As of June 2021, there are more than 100 GWAS studies trying to identify possible candidate genes and human genetic variants that are likely involved in COVID-19 pathogenesis. The main aim of our study was an analysis of common variants (MAF>0.05) that showed evidence of association with COVID-19 in studies and characterization of population variability from data generated by NIPT. Allele frequencies of risk COVID-19

variants from studies identified in the Slovak population were compared with allele frequencies of risk COVID-19 variants in 6 worldwide populations. While previous studies have demonstrated the role of the *ACE2* gene, encoding angiotensin-converting enzyme 2, in host defense against COVID-19 [35–37], our study aimed to analyze population allele frequencies and describe the clinical impacts of relevant variants located in the *ACE2* gene. To our knowledge, this was the first population study of COVID-19 using NIPT data conducted exclusively in the Slovak population.

2 Methods

2.1 Data source

The laboratory procedure used, to generate the NIPT data, were as follows: DNA from plasma of peripheral maternal blood was isolated for NIPT analysis from 1,501 pregnant women after obtaining a written informed consent consistent with the Helsinki declaration from the subjects. The population cohort consisted from women in reproductive age between 17-48 years with a median of 35 years. Genomic information from a sample consisted of maternal and fetal DNA fragments. Each included individual agreed to use their genomic data in an anonymized form for general biomedical research. The NIPT study (study ID 35900_2015) was approved by the Ethical Committee of the Bratislava Self-Governing Region (Sabinovska ul.16, 820 05 Bratislava) on 30th April of 2015 under the decision ID 03899_2015. Blood samples were collected to EDTA tubes and plasma was separated in dual centrifugation procedure. DNA was isolated from 700 μl of plasma using DNA Blood Mini kit (Qiagen, Hilden, DE) according to standard protocol. Sequencing libraries were prepared from each sample using TruSeq Nano kit HT (Illumina, San Diego, CA, USA) following standard protocol with omission of DNA fragmentation step. Individual barcode labelled libraries were pooled and sequenced using low-coverage whole-genome sequencing on an Illumina NextSeq500 platform (Illumina, San Diego, CA, USA) by performing paired end sequencing of 2×35 bases [38].

2.2 Data analysis

The detailed information about mapping, exclusion of overlapping reads, quality control and filtering, realignment, genomic coverage and variant calling is fully described in our previous study, in the section Methods and Results [33]. The datasets generated and analyzed during the current study are available in the DSpace repository, https://dspace.uniba.sk/xmlui/handle/123456789/27.

2.3 Analyses of common variants previously reported to be risk variants for COVID-19

We have performed a literature search and combined genotype data from all previously published studies available online (https://pubmed.ncbi.nlm.nih.gov/) for the years 2020-2021 with key words “COVID-19” and “GWAS”, specifically 114 studies focused on genetic variants associated with COVID-19 disease. Using data from these datasets, we have summarized 29 COVID-19 risk variants, which were then merged with our data of identified variants from NIPT. Risk variants that were not found in NIPT data were excluded from the analysis. All identified variants in the Slovak population used for further analyses were common (MAF>0.05). Subsequently, allele frequencies of COVID-19 risk variants for each population (East Asian, South Asian, African, American, Finnish European and non-Finnish European) were extracted from the gnomAD database available online (v3.0, downloaded from https://gnomad.broadinstitute.org/downloads) and compared with our frequencies determined for the Slovak population from NIPT data. Allele frequency in each population and allele frequency differences were plotted using boxplots. Outliers of boxplots that represent variants with highly different frequencies were annotated via published literature (in dbSNP (https://www.ncbi.nlm.nih.gov/snp/). To assess the relations between allele frequency of COVID-19 risk variants in each population, we also used Principal Component Analysis (PCA) using matplotlib.pyplot library, which reduces the dimension of the data to a graphically interpretable 2D or 3D dimension. Consequently, we obtained information on which populations have similar or different allele frequencies of the identified COVID-19 risk variants.

2.4 Analyses of variants located in *ACE2* gene

While previous studies have demonstrated a possible important role of the *ACE2* gene in COVID-19 infection, we also focused on the study of variants located in the *ACE2* gene in our dataset of NIPT data. First, we filtered out a group of variants located in this gene. The genomic locations of the gene were determined by the GeneCards database (https://www.genecards.org/). Allele frequencies of identified variants for each population (East Asian, South Asian, African, American, Finnish European and non-Finnish European) were extracted from the gnomAD database (v3.0, downloaded from https://gnomad.broadinstitute.org/downloads) and compared with frequencies of variants located in *ACE2* gene determined for the Slovak population from NIPT data. Allele frequency in each population and allele frequency differences were plotted using boxplots. Outliers of boxplots were annotated via published literature and studies (in dbSNP (https://www.ncbi.nlm.nih.gov/snp/).

3 Results

3.1 Analyses of common variants previously reported to be risk variants for COVID-19

We found 29 risk genetic variants associated with COVID-19 disease in literature search of studies available online in PubMed (Table 1).

After merging all identified variants from GWAS (29 risk variants) with our NIPT data, we identified 20 common risk COVID-19 variants, while 9 risk variants that were not called in the Slovak population were excluded from further analysis. The allele frequencies of 20 variants identified in our population sample (Slovak population) and the allele frequencies of variants for 6 world populations (East Asian, South Asian, African, American, Finnish European and non-Finnish European) obtained by gnomAD database are shown in graphical comparison by Boxplots (Figure 1) and PCA (Figure 2). The median allele frequency for the Slovak population reached the

value of 0.2712, which is closest to the value of the median of the American population (MED=0.2972). PCA placed our sample set most closely to the non-Finnish European population.

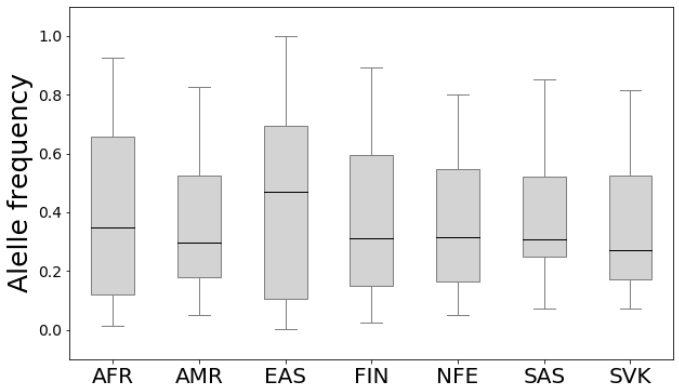

Fig. 1. Boxplots show allele frequency of 20 risk COVID-19 variants identified from GWAS studies for the Slovak and the other six world populations. AFR, African population; AMR, American population; EAS, East Asian population; FIN, European (Finnish) population; NFE, European (non-Finnish) population; SAS, South Asian population; SVK, Slovak population.

Table 1. 29 risk variants associated with COVID-19 in GWAS studies from 2020/2021.

ID_SNP	GENE	PMID
rs2252639	IFNAR2	33259846
rs200008298	DNAH7	33200144
rs183712207	DNAH7	33200144
rs191631470	DNAH7	33200144
rs1024611	CCL2	33133166
rs1800450	MBL	33133166
rs2280788	CCL5	33133166
rs2248690	AHSG	33133166
rs2304237	ICAM3	33133166
rs2430561	IFN	33133166
rs4804803	CD209	33133166
rs2070874	IL4	33133166
rs2070788	TMPRSS2	33133166
rs383510	TMPRSS2	33133166
rs4932178	FURIN	33133166
rs16944	IL1B	33868239
rs2275913	IL17A	33868239
rs1800795	IL6	33868239
rs1800629	TNF	33868239
rs1143633	IL1B	33868239
rs3917332	IL1R1	33868239
rs2232354	IL1RN	33868239
rs1800797	IL6	33868239
rs1800796	IL6	33868239
rs1801274	FCGR2A	33868239
rs11385942	LZTFL1	33868239
rs13078854	LZTFL1	33868239
rs657152	ABO	33868239
rs9411378	ABO	33868239

Next, we compared known allele frequencies of 20 COVID-19 risk variants in our sample set from the Slovak population to allele frequencies of these variants in six world populations. The final findings of allele frequency differences are shown in Figure 3. We identified 2 outliers, rs1801274 and rs383510, in Slovak-Finnish European population comparison and Slovak-non-Finnish European population comparison (Table 2). The rs1801274 is a missense variant located in the *FCGR2A* gene with interpretation in the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/), which aggregates information about genomic variation and its relationship to human health, as drug response. The rs383510 variant is classified as an intron variant in the *TMPRSS2* gene, not reported in ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/).

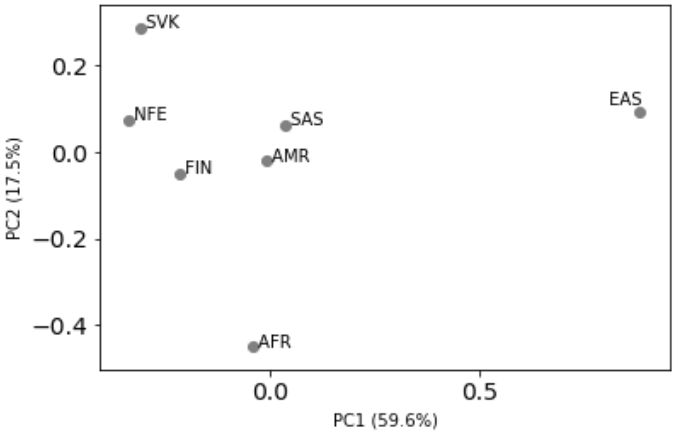

Fig. 2. PCA plot illustrates the allele frequency of 20 risk COVID-19 variants identified from GWAS for the Slovak and the other six world populations.

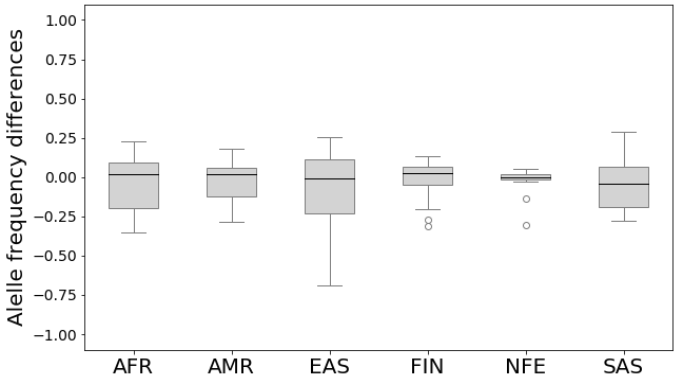

Fig. 3. Boxplots show allele frequency differences of Slovak and the other six world populations for 20 risk COVID-19 variants identified from GWAS. AFR, Slovak-African population; AMR, Slovak-American population; EAS, Slovak-East Asian population; FIN, Slovak-European (Finnish) population; NFE, Slovak-European (non-Finnish) population; SAS, Slovak-South Asian population.

Table 2. Outliers identified in boxplots that show allele frequency differences of Slovak and the other six world populations for 20 risk COVID-19 variants identified from GWAS.

	ALLELE FREQUENCY DIFFERENCES					
SNP	SVK-AFR	SVK-AMR	SVK-EAS	SVK-FIN	SVK-NFE	SVK-SAS
rs1801274	-0.3416	-0.2869	-0.1328	-0.3118	-0.3082	-0.1901
rs383510	-0.2751	-0.2149	-0.2979	-0.2701	-0.1384	-0.2049

.

3.2 Analyses of variants located in *ACE2* gene

In the analysis of variants located in the *ACE2* gene, we identified 79 common variants (MAF > 0.05) in our sample set from NIPT. The allele frequencies of 79 variants identified in our population sample (Slovak population) and the allele frequencies of these variants for 6 world populations (East Asian, South Asian, African, American, Finnish European and non-Finnish European) obtained by gnomAD database are shown in graphical comparison by Boxplots (Figure 4). The median allele frequency for the Slovak population reached the lowest value of 0.5634, which is closest to the value of the median of the non-Finnish European population (MED=0.622636).

In the next step, to identify variants having significantly different frequencies, we compared known allele frequencies of 79 variants located in the *ACE2* gene identified in our sample set from the Slovak population to allele frequencies of these variants in six gnomAD world populations. The final findings of allele frequency differences are shown in Figure 5. By comparing the allele frequency of variants of the Slovak and six world populations, we identified a total of 8 outliers. The variation type of all outliers was "intronic variant" and the clinical significance of all outliers was not reported in ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/).

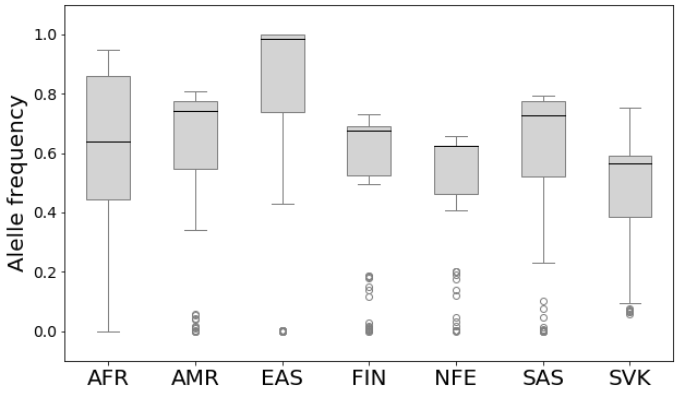

Fig. 4. Boxplots show allele frequency of 79 variants located in the *ACE2* gene for the Slovak and the other six world populations. AFR, African population; AMR, American population; EAS, East Asian population; FIN, European (Finnish) population; NFE, European (non-Finnish) population; SAS, South Asian population; SVK, Slovak population.

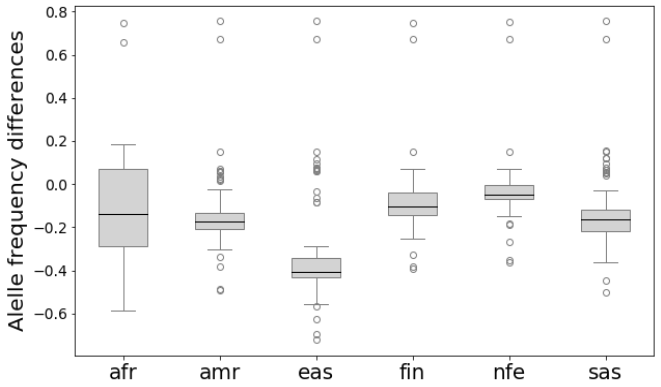

Fig. 5. Boxplots show differences of Slovak and the other six world populations in allele frequency for 79 variants located in the *ACE2* gene. AFR, Slovak-African population; AMR, Slovak-American population; EAS, Slovak-East Asian population; FIN, Slovak-European (Finnish) population; NFE, Slovak-European (non-Finnish) population; SAS, Slovak-South Asian population.

Table 3. Outliers identified in boxplots that show differences of Slovak and the other six world populations in allele frequency for 79 variants located in the *ACE2* gene.

	ALLELE FREQUENCY DIFFERENCES					
SNP	SVK-AFR	SVK-AMR	SVK-EAS	SVK-FIN	SVK-NFE	SVK-SAS
rs1849701	-0.2579	-0.4862	-0.7208	-0.3828	-0.3505	-0.4449
rs869212298	-0.0101	-0.2998	-0.5512	-0.2527	-0.1906	-0.2793
rs35972066	-0.1318	-0.3815	-0.6282	-0.3301	-0.2670	-0.3635
rs34730726	-0.5851	-0.4906	-0.6971	-0.3936	-0.3634	-0.4999
rs67054961	0.6549	0.6723	0.6723	0.6723	0.6697	0.6723
rs111691073	0.7471	0.7544	0.7544	0.7489	0.7511	0.7544
rs2316904	-0.4195	-0.3384	-0.5588	-0.2404	-0.1844	-0.3429
rs114606371	0.1472	0.1474	0.1474	0.1474	0.1474	0.1474

4 Discussion

We performed a literature overview of studies from 2020-2021 that included genetic variants reported to be associated with COVID-19 susceptibility and/or severity and others implicated in the biological pathway of the COVID-19 disease. We compared the allele frequencies of identified variants between Slovak and 6 worldwide populations and, in addition, we also focused on variants located in the *ACE2* gene. To our knowledge, the present study is the first population analysis of COVID-19 variants worldwide and also in the Slovak population using NIPT data. We illustrate the utility of these genomic data for clinical genetics and population studies.

By pooling data of risk variants associated with COVID-19 and data variants in our population sample from NIPT, we have identified 20 common risk variants (MAF> 0.05). When we compared allele frequencies of these variants to allele frequencies in six gnomAD world populations, finally 2 variants were found that showed statistically significant differences in population allele frequencies - rs383510 and rs1801274.

The first intronic SNP, rs383510, is located in the gene *TMPRSS2* frequently discussed in the COVID-19 studies. Together with the *ACE2* gene, the gene *TMPRSS2* is the main host cell entry factor critical for SARS-CoV-2 infection. The spike (S) glycoprotein of the virus binds to the *ACE2* making it essential for the entry of the virus into the host cell [37,39,40]. In addition, S-protein priming by the serine protease *TMPRSS2* allows the fusion of viral and cellular membranes, resulting in virus entry and replication in the host cells [41]. Cheng et al. previously reported that the rs383510 variant, situated in the putative regulatory region with enhancer activity, is significantly associated with the susceptibility to influenzas such as A(H7N9) and A(H1N1)pdm09 influenza [42]. Another study confirmed the effect of rs383510 on the expression of *TMPRSS2* in lung tissues, while the frequencies of variant alleles vary between populations. The rs383510 TT genotype was associated with higher expression of *TMPRSS2* in the lung compared to the CT

and CC genotypes. In addition, the frequency of rs383510 T allele is lower in East Asian populations compared to European and American populations suggesting that a relatively high percentage of Europeans and Americans may have upregulated *TMPRSS2* expression [43]. In our study, the frequency of rs383510 (T/C) was the lowest in Slovak population and the highest frequency was identified in the East Asian population. Schönlfelder et al. analyzed the association of the rs3835510 variant with susceptibility to SARS-CoV-2 infection and severity of COVID-19 in 239 SARS-CoV-2-positive and 253 SARS-CoV-2-negative patients. The results showed that the CC genotype of *TMPRSS2* rs383510 was associated with a 1.73-fold increased SARS-CoV-2 infection risk in a German cohort [44].

The second identified variant in the *FCGR2A* gene, rs1801274, is a missense variant leading to an amino-acid substitution of histidine by arginine at position 131 (H131R). This variant was significantly associated with the risk of severe pneumonia in A/H1N1 influenza infection, bacteremic pneumococcal pneumonia infection, and the severity of community-acquired pneumonia [45]. Another meta-analysis demonstrated that the rs1801274 polymorphism was associated with the susceptibility to multiple autoimmune diseases, including Kawasaki disease and Ulcerative colitis. It has also been reported that the rs1801274 polymorphism may be associated with susceptibility to multiple autoimmune diseases in the Asia population [46].

Regarding genetic variants located in the *ACE2* gene, we did not observe any association between identified variants and SARS-CoV-2 infection. On the one hand, the sample size was relatively small, and it is strongly biased towards the healthy population of females. On the other hand, this could be due to the biological function of *ACE2* as a mediator of cell entry for SARS-CoV-2, so further analyses are needed to conclusively clarify the influence of *ACE2* variants on disease severity.

As the COVID-19 pandemic creates a global crisis and has already had a serious impact on the world, it is crucial to determine how host genetic factors link to clinical outcomes. Several GWAS studies have focused on discovering the influence of host genetic factors on SARS-CoV-2 infection risk or COVID-19 severity. Nevertheless, with the information available to date, not everything has been resolved about the genetic involvement in COVID-19 susceptibility or severity, and new knowledge in the field is continuously generated. Since NIPT expands rapidly to millions of individuals each year, the reuse of these data reduces the cost of large-scale population studies and likely provides an acceptable background for information about genomic variation.

Funding

This publication is the result of support from the Operational Programme Integrated Infrastructure for the project: Pangenomics for personalized clinical management of infected persons based on identified viral genome and human exoma (Code ITMS:313011ATL7), co-financed by the European Regional Development Fund.

Conflict of interest

All authors declare that there is no conflict of interest and they have seen and approved the manuscript submitted.

References

[1] Driggin E, Madhavan MV, Bikdeli B, Chuich T, Laracy J, Biondi-Zoccai G, et al., "Cardiovascular Considerations for Patients, Health Care Workers, and Health Systems During the COVID-19 Pandemic," *J Am Coll Cardiol.* 2020;75: 2352–2371.

[2] Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ, et al., "COVID-19: consider cytokine storm syndromes and immunosuppression," *Lancet.* 2020;395: 1033–1034.

[3] Terpos E, Ntanasis-Stathopoulos I, Elalamy I, Kastritis E, Sergentanis TN, Politou M, et al., "Hematological findings and complications of COVID-19," *Am J Hematol.* 2020;95: 834–847.

[4] Fricke-Galindo I, Falfán-Valencia R., "Genetics Insight for COVID-19 Susceptibility and Severity: A Review," *Front Immunol.* 2021;12: 622176.

[5] Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. "Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding," *Lancet.* 2020;395: 565–574.

[6] Majumder J, Minko T. "Recent Developments on Therapeutic and Diagnostic Approaches for COVID-19," *AAPS J.* 2021;23: 14.

[7] Guarner J. "Three Emerging Coronaviruses in Two Decades," *American Journal of Clinical Pathology.* 2020. pp. 420–421.

[8] Rajgor DD, Lee MH, Archuleta S, Bagdasarian N, Quek SC. "The many estimates of the COVID-19 case fatality rate," *Lancet Infect Dis.* 2020;20: 776–777.

[9] Sun Q, Qiu H, Huang M, Yang Y. "Lower mortality of COVID-19 by early recognition and intervention: experience from Jiangsu Province," *Ann Intensive Care.* 2020;10: 33.

[10]Vincent J-L, Taccone FS. "Understanding pathways to death in patients with COVID-19," *The Lancet. Respiratory medicine.* 2020. pp. 430–432.

[11] Grasselli G, Zangrillo A, Zanella A, Antonelli M, Cabrini L, Castelli A, et al. "Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy," *JAMA.* 2020;323: 1574–1581.

[12] Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. "Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area," *JAMA.* 2020. p. 2052.

[13] Xu X-W, Wu X-X, Jiang X-G, Xu K-J, Ying L-J, Ma C-L, et al. "Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-Cov-2) outside of Wuhan, China: retrospective case series," *BMJ.* 2020;368: m606.

[14] Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. “Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study,” *Lancet.* 2020;395: 507–513.

[15] Wu Z, McGoogan JM. “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*. 2020;323: 1239–1242.

[16] Han C, Duan C, Zhang S, Spiegel B, Shi H, Wang W, et al. “Digestive Symptoms in COVID-19 Patients With Mild Disease Severity: Clinical Presentation, Stool Viral RNA Testing, and Outcomes,” *Am J Gastroenterol.* 2020;115: 916–923.

[17] Tian S, Hu N, Lou J, Chen K, Kang X, Xiang Z, et al. “Characteristics of COVID-19 infection in Beijing,” *J Infect.* 2020;80: 401–406.

[18] Singhal T. “A Review of Coronavirus Disease-2019 (COVID-19),” *Indian J Pediatr*. 2020;87: 281–286.

[19] Pan L, Mu M, Yang P, Sun Y, Wang R, Yan J, et al. “Clinical Characteristics of COVID-19 Patients With Digestive Symptoms in Hubei, China: A Descriptive, Cross-Sectional, Multicenter Study,” *Am J Gastroenterol*. 2020;115: 766–773.

[20] Cantalupo S, Lasorsa VA, Russo R, Andolfo I, D’Alterio G, Rosato BE, et al. “Regulatory Noncoding and Predicted Pathogenic Coding Variants of Predispose to Severe COVID-19,” *Int J Mol Sci.* 2021;22.

[21] Liu X, Zhou H, Zhou Y, Wu X, Zhao Y, Lu Y, et al. “Risk factors associated with disease severity and length of hospital stay in COVID-19 patients,” *The Journal of infection*. 2020. pp. e95–e97.

[22] Scully EP, Haverfield J, Ursin RL, Tannenbaum C, Klein SL. “Considering how biological sex impacts immune responses and COVID-19 outcomes,” *Nat Rev Immunol*. 2020;20: 442–447.

[23] Gebhard C, Regitz-Zagrosek V, Neuhauser HK, Morgan R, Klein SL. “Impact of sex and gender on COVID-19 outcomes in Europe,” *Biology of Sex Differences*. 2020.

[24] Jutzeler CR, Bourguignon L, Weis CV, Tong B, Wong C, Rieck B, et al. “Comorbidities, clinical signs and symptoms, laboratory findings, imaging features, treatment strategies, and outcomes in adult and pediatric patients with COVID-19: A systematic review and meta-analysis,” *Travel Med Infect Dis*. 2020;37: 101825.

[25] Guan W-J, Ni Z-Y, Hu Y, Liang W-H, Ou C-Q, He J-X, et al. “Clinical Characteristics of Coronavirus Disease 2019 in China,” *N Engl J Med*. 2020;382: 1708–1720.

[26] Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. “Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study,” *The Lancet*. 2020. pp. 1054–1062.

[27] Cummings MJ, Baldwin MR, Abrams D, Jacobson SD, Meyer BJ, Balough EM, et al. “Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study,” *medRxiv. 2020*.

[28] Hu J, Li C, Wang S, Li T, Zhang H. “Genetic variants are identified to increase risk of COVID-19 related mortality from UK Biobank data,” *medRxiv*. 2020.

[29] Anastassopoulou C, Gkizarioti Z, Patrinos GP, Tsakris A. “Human genetic factors associated with susceptibility to SARS-CoV-2 infection and COVID-19 disease severity,” *Hum Genomics*. 2020;14: 40.

[30] Yildirim Z, Sahin OS, Yazar S, Bozok Cetintas V. “Genetic and epigenetic factors associated with increased severity of Covid-19,” *Cell Biol Int.* 2021;45: 1158–1174.

[31] Severe Covid-19 GWAS Group, Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, et al. “Genomewide Association Study of Severe Covid-19 with Respiratory Failure,” *N Engl J Med.* 2020;383: 1522–1534.

[32] Schmiedel BJ, Chandra V, Rocha J, Gonzalez-Colin C, Bhattacharyya S, Madrigal A, et al. “COVID-19 genetic risk variants are associated with expression of multiple genes in diverse immune cell types,” *bioRxiv*. 2020.

[33]Budis J, Gazdarica J, Radvanszky J, Harsanyova M, Gazdaricova I, Strieskova L, et al. “Non-invasive prenatal testing as a valuable source of population specific allelic frequencies,” *J Biotechnol.* 2019;299: 72–78.

[34] Pös O, Budis J, Kubiritova Z, Kucharik M, Duris F, Radvanszky J, et al. “Identification of Structural Variation from NGS-Based Non-Invasive Prenatal Testing,” *Int J Mol Sci*. 2019;20.

[35] Esler M, Esler D. “Can angiotensin receptor-blocking drugs perhaps be harmful in the COVID-19 pandemic?,” *J Hypertens.* 2020;38: 781–782.

[36] Gemmati D, Tisato V. “Genetic Hypothesis and Pharmacogenetics Side of Renin-Angiotensin-System in COVID-19. Genes,” . 2020;11.

[37] Benetti E, Tita R, Spiga O, Ciolfi A, Birolo G, Bruselles A, et al. “ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population,” *Eur J Hum Genet.* 2020;28: 1602–1614.

[38] Minarik G, Repiska G, Hyblova M, Nagyova E, Soltys K, Budis J, et al. “Utilization of Benchtop Next Generation Sequencing Platforms Ion Torrent PGM and MiSeq in Noninvasive Prenatal Testing for Chromosome 21 Trisomy and Testing of Impact of In Silico and Physical Size Selection on Its Analytical Performance,” *PLoS One*. 2015;10: e0144811.

[39] Asselta R, Paraboschi EM, Mantovani A, Duga S. “ACE2 and TMPRSS2 variants and expression as candidates to sex and country differences in COVID-19 severity in Italy,” *Aging*. 2020. pp. 10087–10098.

[40] Singh H, Choudhari R, Nema V, Khan AA. “ACE2 and TMPRSS2 polymorphisms in various diseases with special reference to its impact on COVID-19 disease,” *Microb Pathog*. 2021;150: 104621.

[41] Andolfo I, Russo R, Lasorsa VA, Cantalupo S, Rosato BE, Bonfiglio F, et al. “Common variants at 21q22.3 locus influence and gene expression and susceptibility to severe COVID-19,” *iScience*. 2021;24: 102322.

[42] Cheng Z, Zhou J, To KK-W, Chu H, Li C, Wang D, et al. “Identification of TMPRSS2 as a Susceptibility Gene for Severe 2009 Pandemic A(H1N1) Influenza and A(H7N9) Influenza,” *J Infect Dis.* 2015;212: 1214–1221.

[43] Irham LM, Chou W-H, Calkins MJ, Adikusuma W, Hsieh S-L, Chang W-C. “Genetic variants that influence SARS-CoV-2 receptor TMPRSS2 expression among population

cohorts from multiple continents," *Biochem Biophys Res Commun.* 2020;529: 263–269.
[44] Schönfelder K, Breuckmann K, Elsner C, Dittmer U, Fistera D, Herbstreit F, et al. "Polymorphisms and Susceptibility to Severe Acute Respiratory Syndrome Coronavirus Type 2 Infection: A German Case-Control Study," *Front Genet.* 2021;12: 667231.
[45] Shi X, Ma Y, Li H, Yu H. "Association between FCGR2A rs1801274 and MUC5B rs35705950 variations and pneumonia susceptibility," *BMC Med Genet.* 2020;21: 71.
[46] Zhang C 'e, Wang W, Zhang H 'e, Wei L, Guo S. "Association of FCGR2A rs1801274 polymorphism with susceptibility to autoimmune diseases: A meta-analysis," *Oncotarget.* 2016;7: 39436–39443.