Received: 23 April 2021 | Revised: 18 July 2021 | Accepted: 22 July 2021

DOI: 10.1002/rmv.2283

REVIEW

WILEY

Systematic review of host genetic association with Covid-19 prognosis and susceptibility: What have we learned in 2020?

João Locke Ferreira de Araújo[1,2,3] | Diego Menezes[1,2,3] | Julia Maria Saraiva-Duarte[1,3] | Luciana de Lima Ferreira[1,2,3] | Renato Santana de Aguiar[1,2,3] | Renan Pedra de Souza[1,2,3]

[1]Laboratório de Biologia Integrativa, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

[2]Grupo de Pesquisa em Bioestatística e Epidemiologia molecular, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

[3]Programa de Pós-graduação em Genética, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

Correspondence

Renan Pedra de Souza, Laboratório de Biologia Integrativa, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas; Universidade Federal de Minas Gerais, Av. Antonio Carlos, 6627 ICB, Pampulha 31270901, Belo Horizonte, Minas Gerais, Brazil.
Email: renanrps@ufmg.br

Funding information

Fundação de Amparo à Pesquisa do Estado de Minas Gerais

Summary

Biomarker identification may provide strategic opportunities to understand disease pathophysiology, predict outcomes, improve human health, and reduce healthcare costs. The highly heterogeneous Covid-19 clinical manifestation suggests a complex interaction of several different human, viral and environmental factors. Here, we systematically reviewed genetic association studies evaluating Covid-19 severity or susceptibility to SARS-CoV-2 infection following PRISMA recommendations. Our research comprised papers published until December 31st, 2020, in PubMed and BioRXiv databases focusing on genetic association studies with Covid-19 prognosis or susceptibility. We found 20 eligible genetic association studies, of which 11 assessed Covid-19 outcome and 14 evaluated infection susceptibility (five analyzed both effects). Q-genie assessment indicated moderate quality. Five large-scale association studies (GWAS, whole-genome, or exome sequencing) were reported with no consistent replication to date. Promising hits were found on the 3p21.31 region and ABO locus. Candidate gene studies examined *ACE1*, *ACE2*, *TMPRSS2*, *IFITM3*, *APOE*, *Furin*, *IFNL3*, *IFNL4*, *HLA*, *TNF-a* genes, and ABO system. The most evaluated single locus was the *ABO*, and the most sampled region was the *HLA* with three and five candidate gene studies, respectively. Meta-analysis could not be performed. Available data showed the need for further reports to replicate claimed associations.

KEYWORDS

candidate genetic variants, Covid-19, genetic association, polymorphisms, SARS-CoV-2, susceptibility

Abbreviations: ABO, ABO blood group system; ACE1, angiotensin-converting enzyme-1; ACE2, angiotensin-converting enzyme-2; APOE, apolipoprotein E; CLUAP1, Clusterin Associated Protein 1; Covid-19, Coronavirus disease; DES, Desmin; DNAH7, Dynein Axonemal Heavy Chain 7; GOLGA8B, Golgin A8 Family Member B; GWAS, Genome-wide association study; HLA, Human leukocyte antigen; IFITM3, Interferon Induced Transmembrane Protein 3; IFNL3, Interferon Lambda 3; IFNL4, Interferon Lambda 4; IRF7, Interferon Regulatory Factor 7; MUC2, Mucin 2; PCDH15, Protocadherin Related 15; PCR, Polymerase chain reaction; RIMBP3, RIMS binding protein 3; RT, reverse transcriptase; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SPEG, Striated Muscle Enriched Protein Kinase; SSP, sequence-specific oligonucleotide; STREGA, Strengthening the reporting of genetic association studies; STXBP5, Syntaxin Binding Protein 5; TLR3, Toll-Like Receptor 3; TMEM189, Transmembrane protein 189; TMPRSS2, Transmembrane protease, serine 2; TNF-α, Tumor necrosis factor-alpha; TOMM7, Translocase of Outer Mitochondrial Membrane 7; UBE2V1, Ubiquitin Conjugating Enzyme E2 V1; WSB1, WD Repeat and SOCS Box Containing 1.

1 | INTRODUCTION

Coronavirus disease (Covid-19) pandemic remains overwhelming healthcare systems and damaging economies. People infected with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) present a wide range of severity of illness, from asymptomatic or mild to severe disease and death. Recent results estimate over 20.5 million life-years have been lost due to Covid-19 globally.[1] The highly heterogeneous Covid-19 clinical manifestations suggest a complex interaction of several different human, viral and environmental factors playing a role in Covid-19 prognosis.[2-4] Understanding mechanisms leading to severe cases is of great importance for therapeutic development and pandemic control. Furthermore, infection susceptibility has been associated with several factors.[5,6]

As to social-environmental aspects, the pandemic exposed pre-existing health and social differences between historically vulnerable populations. A remarkable contrast between the mortality rate from Covid-19 in minority groups exists compared to privileged social stratum.[7] As to the pathogen aspects, the SARS-CoV-2 genome has almost 30,000 base pairs with structural genes (spike, nucleocapsid, membrane, and envelope) and non-structural proteins (involved in replication).[8] SARS-CoV-2 genome research has demonstrated viral diversity[9,10] may be related to pathogenicity, transmissibility, and, more recently, mortality.[11] Variability on the *S* viral gene seems relevant since it codes for the spike protein that interacts with two crucial cell entry factors: the human angiotensin-converting enzyme-2 (ACE2) receptor and the cellular serine protease TMPRSS2. Recent results indicate other possible human targets (e.g., cathepsin L).[12]

As to the host aspects, structural data analysis proposed that *ACE2* gene variants can alter host-virus interaction and Covid-19 susceptibility.[13] Apart from ACE2, several other proteins have been associated with Covid-19 pathogenesis and immune response. Immunomodulatory molecules seem to play a crucial role (e.g., cytokine storm). It would be possible to hypothesize that polymorphisms in their genes could contribute to Covid-19 prognosis.[14] Here, we systematically reviewed genetic association studies evaluating Covid-19 severity or susceptibility to SARS-CoV-2 infection.

2 | METHODS

2.1 | Systematic review

We registered a study protocol on PROSPERO (CRD42020187270). Preferred Reporting Items for Systematic reviews and Meta-analysis (PRISMA) was adopted as a guideline for reporting this systematic review.[15] Study selection was carried out in three phases: identification, screening, and eligibility. Identification was performed by searching on two databases: PubMed and BioRXiv. The bibliographic search included all studies published untill 31 December 2020, using the search arguments listed in the supplementary material (supplementary material I).

Two independent researchers conducted the screening of the articles. Inclusion criteria were primary articles covering human genetics association with Covid-19 susceptibility and/or prognosis, while exclusion criteria were review articles or primary articles not covering genetic association with Covid-19 susceptibility and/or prognosis. A systematic review flowchart was prepared following PRISMA specifications.

2.2 | Article quality analyses

We assessed study quality using the Q-Genie tool[16] performed by two independent researchers. This instrument contains 11 questions to be marked on a seven-point Likert scale examing several aspects of a genetic association study: scientific basis for the development of the research question, ascertainment of comparison groups (e.g., cases and controls), technical and non-technical classification of tested genetic variants (e.g., genotyping call rates, blinded experiments), classification of the outcome (e.g., sampling strategy, definition criteria), discussion of sources of bias, appropriateness of sample size, description of planned statistical analyses, statistical methods applied, test of assumptions in the genetic studies (e.g., Hardy–Weinberg equilibrium) and appropriate interpretation of the results.[16,17] Since all studies used a case-control design, cut-offs were $\leq$35 for poor, >35 for moderate, and $\leq$45 for good quality, according to Sohabi et al.[17] (total sum may vary from 7 to 77 points).

3 | RESULTS

Our literature search returned 1633 records from the two databases (Figure 1). Three additional articles were added from other sources (e.g., cross-referencing), leading to 1636 records. We excluded 1587 articles after reading titles and abstracts. We removed another 29 manuscripts following full-text analysis (supplementary material II). In the end, 20 studies were eligible for the qualitative synthesis.

We found 11 studies addressing genetic influence on the prognosis of Covid-19 (Table 1) and 14 studies exploring the susceptibility to Sars-CoV-2 infection (Table 2). Five studies worked with both approaches.[18-22] Two studies proposed to work with prognosis, but the outcome was susceptibility.[23,24] Study quality assessment resulted in six studies with poor quality, seven moderate, and seven classified as good. The mean quality score reached moderate classification (mean 41.56; standard deviation 9.05). One of the most valuable pieces of information from the Q-Genie usage is evaluating quality dimensions across studies, thus identifying systematic issues. We consistently observed inadequate description of the genotyping process leading Q-Genie item number five to have the lowest mean score. We report that most studies failed to inform whether researchers performed genotyping blinded from case-control information or whether any randomization occurred across cases and controls to avoid batch effects. On the other hand, we found that the most successful quality aspect was presenting the rationale to

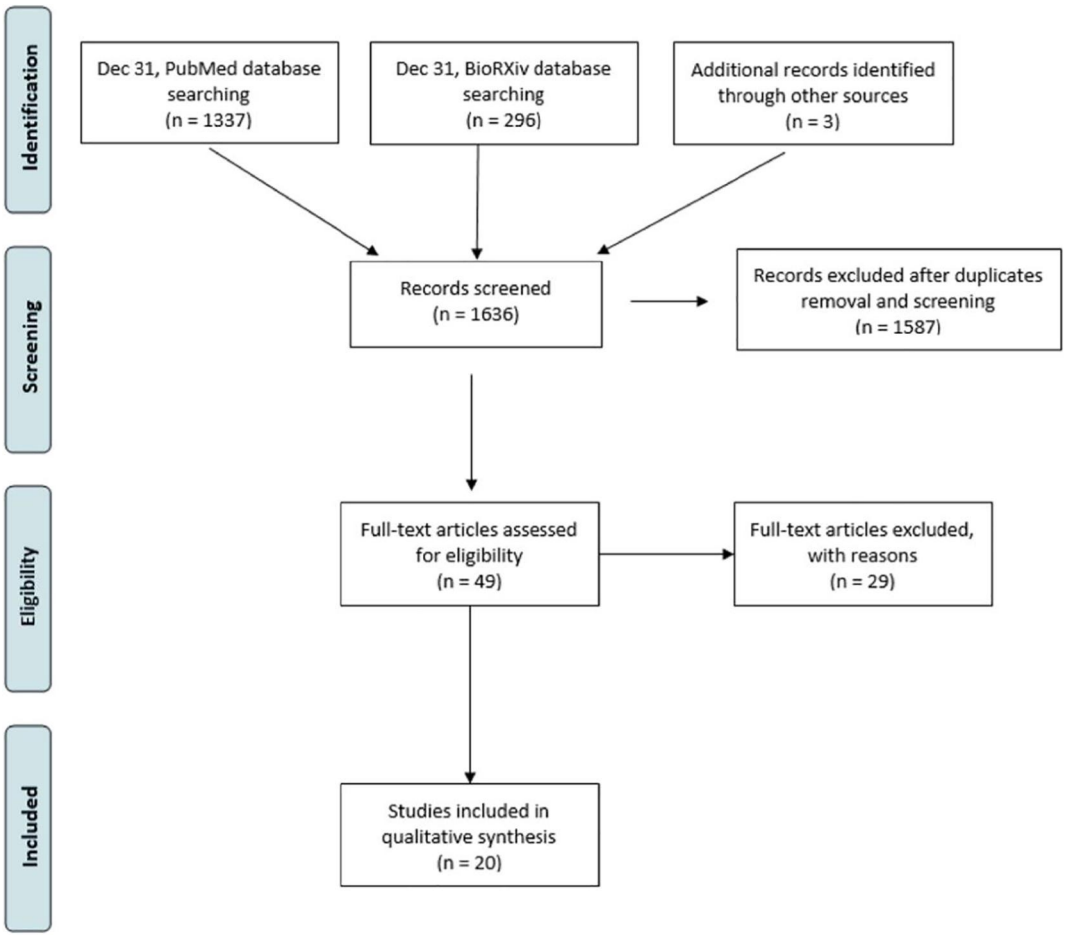

FIGURE 1 PRISMA flow diagram

conduct a genetic association study (Q-Genie item number one–rationale for analysis).

Evidence of genetic association was reported in six of the 11 studies addressing the Covid-19 clinical outcome.[19-21,25-28] Three of these studies were large-scale association with either whole-genome sequencing or GWAS approach.[21,26,27] Hu et al.[27] reported GWAS significant signals on the *DNAH7 CLUAP1*, *DES*, *SPEG*, *STXBP5*, *TOMM7*, *PCDH15*, and *WSB1* genes. Zhang et al.[26] focused on rare variants associated with a monogenic contribution to life-threatening Covid-19 and found 10 variants in the TLR3 and IRF7 genetic pathways. Still, their results were not replicated by Povysil et al.[29] Wang et al. found an association with severity on the *TMEM189–UBE2V1* gene locus (rs6020298) using whole-genome sequencing.[21] Another three candidate gene studies indicated statistically significant loci, each evaluating genes related to immune response (rs12252 *IFITM3*[25]; multiples alleles of *HLA-A*; *HLA-B*; *HLA-C*; *HLA-DRB1*[19]) and ACE2 expression (rs4646994 *ACE1*[20]).

Nine of the 14 studies investigating genetic association with susceptibility found significant evidence of increased risk in several loci [21,23,25,26,31,33–36]. Using whole-genome sequencing, Wang et al.[21] indicated a possible contribution of rs200975425 located in the *GOLGA8B* gene, rs200584390 in *RIMBP3*, and a novel missense variant found in *MUC2*. Using the GWAS approach, Ellinhaus et al. found a hit on chromosome 3p21.31 region.[23] Studies with candidate gene approach suggested roles for several alleles on *HLA* region (C*07:29, B*15:27, B*27:07, DRB1*15:01, DQB1*06:02, C*06:02 and DRB1*07:01 in Novelli et al.[30]; HLA-C*07:29 and HLA-B*15:27 in Wang et al.,[28] and HLA-C*04:01 in Littera et al.[19]) and in genes associated with the viral cell cycle (rs61735794 and rs61735792 located in *TMPRSS2*[31] and *APOE* allele e4[26]). Genetic variance of the ABO blood system was also significant: while A-type subjected showed increased susceptibility, O-type individuals were less likely to be infected.[32,33]

4 | DISCUSSION

Biomarker identification may provide a strategic opportunity to understand disease pathophysiology and predict outcomes, therefore improving human health and reducing healthcare costs. Thus far, the most promising prognosis predictors are age,[34] sex,[35] comorbidities,[36,37] and viral load at the moment of infection.[38] Host genetic variants have been suggested as prognostic and infection susceptibility markers in other infectious diseases, for example, *CCR-5* delta

TABLE 1 Genetic contribution to Covid-19 prognosis. Subtitle: Studies that addressed Covid-19 outcomes found in the database search

Author, year	Country	Sample description	Total sample (n)	Severity (n)	Genotyping	Genes/variants	Results
Candidate gene association							
Zhang et al., 2020[25]	China	Confirmed Covid-19 Patients from Youan Hospital, Beijing	80	Mild (56); Severe (24)	Sanger sequencing	*IFITM3* (rs12252)	Increased severity for CC genotype carriers ($p = 0.00093$; OR = 6.37)
Novelli et al., 2020[53]	Italy	Confirmed Covid-19 patients from Tor Vergata University Hospital (89) and Bambino Gesù Children's Hospital (42), Rome	131	Asymptomatic (17); mild (16); moderate (43); severe (55)	Whole exome sequencing	*ACE2* (rs140312271, rs2285666 and rs41303171)	No association
Gómez et al., 2020[20]	Spain	Confirmed Covid-19 patients from the region of Asturias	204	Mild (137); severe (67)	PCR and PCR-RFLP	*ACE1* (rs4646994); *ACE2* (rs2285666)	Increased severity for ACE1- D carriers (total patients $p = 0.049$, and male patients: $p = 0.043$)
Lorente et al., 2020[18]	Spain	Confirmed Covid-19 patients from 8 Intensive care Units from 6 hospitals of canary Islands	72	Death (10); Survival (62)	PCR-SSP	*HLA-A* (*11); *HLA-C* (*01); *HLA-DQB1* (*04)	No association
Amodio et al., 2020[54]	Italy	Confirmed Covid-19 patients from University Hospital "P. Giaccone" of Palermo, western Sicily	381	Death (32); Intensive care hospitalization (21); Hospitalization (93); Home isolation (235)	PCR-SSP	*IFNL3* (rs12979860); *INFL4* (rs368234815)	No association
Rosenbaum et al., 2020[22]	Several	Alleged (23) and confirmed (18) Covid-19 patients affected by spondyloarthritis from 65 countries	41	10-level scale, being 1 extremely mild symptoms and 10 life-threatening symptoms: Level 1 (1); level 2 (2); level 3 (5); level 4 (4); level 5 (7); level 6 (6); level 7 (7); level 8 (6); level 9 (2); level 10 (1)	Not reported	*HLA-B* (*27)	No association
Littera et al., 2020[19]	Italy	Covid-19 confirmed patients from SS. Trinità Hospital in cagliari and asymptomatic or paucisymptomatic patients were confined to home quarantine in cagliari.	141	Severe (39); asymptomatic or paucisymptomatic (143)	PCR-SSP and next Generation sequencing	*HLA-A; HLA-B; HLA-C; HLA-DRB1* (multiple alleles)	Decreased severity in HLA-A*23 and HLA-DRB1*08 carriers

TABLE 1 (Continued)

Author, year	Country	Sample description	Total sample (n)	Severity (n)	Genotyping	Genes/variants	Results
Large scale association							
Zhang et al., 2020[26]	Several (COVID human genetic effort)	Confirmed Covid-19 patients	1193	Asymptomatic/mild (534); life-threatening (659)	Whole exome or genome sequencing	13 loci associated with interferon I response pathway	Increased life-threatening associated with 10 variants in TLR3- and IRF7 in a monogenic model
Wang et al., 2020[21]	China	Confirmed Covid-19 patients from Shenzhen Third Hospital	332	Asymptomatic (25), mild (12), moderate (225), severe (53), critically (17)	Whole-genome sequencing	Loci across the whole genome	Increased severity for minor allele carriers of TMEM189–UBE2V1-rs6020298 (OR = 1.2)
Hu et al., 2020[27]	UK	Confirmed Covid-19 patients from UK biobank	1778	Death (445); Survival (1333)	GWAS or next Generation sequencing	Loci across the whole genome	Increased mortality for carriers of variants in the following loci: *STXBP5/STXBP5-AS1* (OR = 2.91); *CPQ* (OR = 1.92); *CLUAP1* (OR = 2.72); *WSB1* (OR = 4.23); *DNAH7/SLC39A10* (OR = 2.55); *DES/SPEG* (OR = 2.73); *TOMM7* (OR = 2.41); *PCDH15* (OR = 2.52)
Povysil et al., 2020[29]	USA; Canada; Saudi Arabian and Qatar	Confirmed Covid-19 patients from four different cohorts	Columbia University COVID-19 biobank cohort (1153); Biobanque Québec COVID-19 cohort (533); Saudi Arabian COVID-19 cohort (307); Qatar genome Program COVID-19 cohort (700)	Columbia University COVID-19 biobank cohort (severe (480) and mild (673)); Biobanque Québec COVID-19 cohort (severe or die (62); mild (128); no hospitalization (30)); Saudi Arabian COVID-19 cohort (severe (148) and mild or asymptomatic (159)); Qatar genome Program COVID-19 cohort (severe (60) and mild or asymptomatic (640))	GWAS or next Generation sequencing	13 loci associated with interferon I response pathway - same as Zhang et al.[26]	No association

TABLE 2 Genetic contribution to SARS-CoV-2 infection susceptibility

Author, year	Country	Cases (n)	Controls (n)	Genotyping	Genes/variants	Results
Candidate gene association						
Wang et al., 2020[28]	China	Confirmed Covid-19 patients from Zhejiang (82)	Negative controls (3548) were obtained from previous studies of bone marrow from Zhejiang	Next Generation sequencing (patients) and PCR (control)	*HLA-A; HLA-B* (*15:27 and *40:06); *HLA-C* (*07:29 and *08:01G); *HLA-DRB1* (*04:06 and *12:02); *HLA-DRB3/4/5; HLA-DQA1; HLA-DQB1; HLA-DPA1; HLA-DPB1* (*04:01 and *36:01)	Increased susceptibility for HLA-C*07:29 and HLA-B*15:27 allele carriers
Torre-Fuentes et al., 2020[31]	Spain	Confirmed Covid-19 patients from 23 families affected by MultipleSclerosis (7)	Negative controls from 23 families affected by multiple Sclerosis (113). (Unclear definition)	Whole-exome sequencing	*ACE2 (rs35803318 and rs41303171); TMPRSS2 (rs17854725, rs75603675, rs2298659, rs12329760, rs3787950, rs61735794, rs61735792, rs142750000, rs200291871 and rs141788162); Furin (rs6226, rs753334944, rs16944971, rs73489557, rs6225 and ND (c.1956_1956delG 1)*	Increased susceptibility for minor allele carriers of rs61735794 and rs61735792
Fan et al., 2020[32]	China	Confirmed Covid-19 patients from Zhongnan Hospital of Wuhan University (105)	Negative controls from Zhongnan Hospital of Wuhan University (103). No history of respiratory infections and other infectious diseases	*ABO Blood Typing*	*ABO* (A, B, and O)	Increased susceptibility in A-type (OR = 1.33)
Kuo et al., 2020[24]	England	Confirmed Covid-19 patients affected by dementia or delirium from UK biobank (622)	Negative or not-tested controls affected by dementia or delirium from UK biobank (322.326). PCR negative or not tested	GWAS or next Generation sequencing	*APOE* (e3 and e4)	Increased susceptibility for e4e4 genotype (OR = 2.31)
Gómez et al., 2020[20]	Spain	Confirmed Covid-19 patients from the region of Asturias, Northern Spain (204)	Negative controls from the region of Asturias, Northern Spain (536). Healthy population controls (unclear definition)	PCR and PCR-RFLP	*ACE1* (rs4646994); *ACE2* (rs2285666)	No association

TABLE 2 (Continued)

Author, year	Country	Cases (n)	Controls (n)	Genotyping	Genes/variants	Results
Novelli et al., 2020[30]	Italy	Confirmed Covid-19 patients (99)	Negative controls (1017) were previously typed in the laboratory.	Next-generation sequencing	*HLA-B* (*27:07 and *58:01); *HLA-C* (*06:02); *HLA-DRB1* (*07:01 and *15:01); *HLA-DBQ1 (*06:02)*	Increased susceptibility for HLA B*27:07; DRB1*15:01; DQB1*06:02; C*06:02, and DRB1*07:01 alleles
Zhao et al., 2020[33]	China	Covid-19 confirmed patients from the Jinyintan Hospital in Wuhan, Hubei province, China (1775) and Renmin Hospital of Wuhan University, Hubei province, and Shenzhen Third People's Hospital, Guangdong province, China (398).	Negative controls from Wuhan city (3.694) and Shenzhen city (23.396). Non- covid-19 (unclear definition).	*ABO Blood Typing*	*ABO* (A, B, and O)	Increased susceptibility for A-type (OR = 1.279) and decreased susceptibility for O-type (OR = 0.680)
Lorente et al., 2020[18]	Spain	Confirmed Covid-19 patients from 8 Intensive care Units of 6 hospitals at the canary Islands (72)	Negative controls from canary Islands (3.886). Healthy people (unclear definition)	PCR-SSO	*HLA-A; HLA-B; HLA-C; HLA-DRB1; HLA-DQB1*	No association
Rosenbaum et al., 2020[22]	-	Alleged (23) and confirmed (18) Covid-19 patients affected by spondyloarthritis from 65 countries	Negative controls affected by spondyloarthritis from 65 countries (2.795). (Unclear definition)	Not reported	*HLA-B* (*27)	No association
Benetti et al., 2020[55]	Italy	Confirmed Covid-19 patients from, with the contribution of centers in Italy (131).	Negative controls from Italy (258). Healthy people (unclear definition).	Whole-exome sequencing	*ACE2* (p.(Asn720Asp); p.(Lys26Arg), p.(Gly211Arg), p.(Leu351Val) and p.(Pro389His))	No association
Saleh et al., 2020[56]	Egypt	Confirmed Covid-19 patients from Quarantine Department, Mansoura University Hospital (900)	Health care workers in contact with the patients (184). Health care workers (unclear definition).	PCR	*TNF-a G-308 A*	No association
Littera et al., 2020[19]	Italy	Covid-19 confirmed patients from SS. Trinità Hospital in cagliari (39) and asymptomatic or paucisymptomatic patients were confined to home quarantine (143): (182).	Negative controls from Sardinian (619) RT-PCR negative from a nasopharyngeal swab.	PCR-SSP and next Generation sequencing	*HLA-A; HLA-B; HLA-C; HLA-DRB1* (multiple alleles)	Increased susceptibility for HLA-C*04:01 allele carriers (OR = 1.8)

(Continues)

TABLE 2 (Continued)

Author, year	Country	Cases (n)	Controls (n)	Genotyping	Genes/variants	Results
Large scale association						
Ellinghaus et al., 2020[23]	Italy and Spain	Confirmed Covid-19 patients from 7 hospitals of Milan, Barcelona, Madrid, and San Sebastian (1610)	Negative controls from 7 hospitals of Milan, Barcelona, Madrid, and San Sebastian (2.205). Persons with unknown SARS-CoV-2 infection	GWAS	Loci across the whole genome	Increased susceptibility in 3p21.31 region minor allele (OR = 1.77) and A-type (OR 1.45)
Wang et al., 2020[21]	China	Confirmed Covid-19 patients from Shenzhen Third Hospital (284)	Negative controls from 1000 genome project (301) from the ChineseReference Panel program (665)	Whole-genome sequencing	Loci across the whole genome	Increased susceptibility for minor allele carriers of novel missense variant in *MUC2* (OR = 18), *GOLGA8B* rs200975425 (OR = 5.4) and *RIMBP3* rs200584390 (OR = 9.29)

32 with HIV[39]; *TMPRSS2*,[40] *TLR-3* genes with influenza[41,42]; and blood group with dengue.[43]

Throughout last year, 20 genetic association studies were conducted. The most evaluated single locus was the *ABO*, and the most sampled region was the *HLA* with three[23,32,33] and five[18,19,22,28,30] candidate gene studies, respectively. We did not perform a meta-analysis because there was no replication for the same genetic variant or divergence on phenotype definition. Zhang et al.[26] and Povysil et al.[29] studies were the closest studies regarding experimental design and genetic variants examined, with both aiming to find rare variants associated with disease severity in the interferon I response pathway. They reached divergent results, but different control definitions and confounder variant treatment, such as age, may have contributed.[44] The need for replication studies has been extensively discussed to assess the credibility of the initial association, therefore, avoiding the winner's curse phenomenon.[45] Whenever possible, replication studies should be performed in larger samples and consider bias due to population stratification, misclassification of clinical outcome, among others.

In 2021, large consortia organized last year published highly expected studies. The COVID-19 Host Genetics Initiative[46] presented results from three genome-wide association meta-analyses comprised of up to 49,562 Covid-19 patients from 46 studies across 19 countries.[47] They report 13 genome-wide significant loci. Of particular interest, the 3p21.31 region seems to be associated with infection susceptibility[47,] while Ellinghaus et al.[23] significantly correlated it with severity. The ABO locus also appeared relevant for susceptibility.[47] Similar results were also found in a study conducted by the 23 and Me using their biobank.[48] Another critical large-scale association study was published reporting data from more than half-million subjects, of which 20,952 had Covid-19.[54] They did not found rare variants associated either exome wide or when specifically focusing on (1) 13 interferon pathway genes in which rare deleterious variants have been reported in individuals with severe COVID-19,[29] (2) 281 genes located in susceptibility loci identified by the COVID-19 Host Genetics Initiative,[47] or (3) 32 additional genes of immunologic relevance and/or therapeutic potential.[49] Therefore, recent results also indicate that additional research is needed.

Quality assessment of the included studies points to several interesting questions. Firstly, control group definition varies across studies aiming to perform genetic association with the same outcome. Two reports[23,24] examining Covid-19 prognosis used healthy subjects as controls, while other studies with equivalent phenotype used asymptomatic or mild Covid-19 patients. While we believe healthy individuals would be suitable as a control in susceptibility studies, it would be recommended to assume a good prognosis only in SARS-CoV-2 challenged subjects. In other words, healthy subjects from previously organized biobanks may include patients who will present a worse prognosis when infected, thus biasing the control group. Analysis with asymptomatic or paucisymptomatic individuals could also provide relevant results on the genetic basis related to all Covid-19 manifestations.[50] Secondly, we observed divergences in the clinical or molecular inclusion criteria for negative patients. Some studies

required molecular testing while others didn't, that is, only clinical symptomatology was assessed.[22] It is also relevant to point out that several studies were not transparent regarding their criteria, as indicated by unclear definition in Table 2. Thirdly, most studies lack basic technical information (e.g., blinded genotyping, randomization, or the number of batches in which samples were processed) that may be different sources of relevant bias. A powerful tool to avoid further inadequate reporting of genetic association studies is the "strengthening the reporting of genetic association studies" (STREGA) report.[51] It includes a detailed checklist with elements that should be presented in a genetic association publication. While the STREGA recommendations do not aim to influence how a genetic association study should be designed, it seeks to enhance reporting transparency, thus also improving reproducibility.

While this review has highlighted many genes that may be potentially associated with Covid-19 prognosis and infection susceptibility, limitations such as lack of reproducibility, quality of reporting, and quality of assessment remain a significant concern. Therefore, results should be taken with caution. Future studies are also warranted in underrepresented ancestries since the allelic frequency, and linkage disequilibrium may vary across different populations.[52]

ACKNOWLEGEMENT

None.

AUTHOR CONTRIBUTION

João Locke Ferreira de Araújo, Diego Menezes, Luciana de Lima Ferreira, Renato Santana de Aguiar, and Renan Pedra de Souza wrote the systematic review protocol. João Locke Ferreira de Araújo and Diego Menezes conducted the systematic review. João Locke Ferreira de Araújo and Julia Maria Saraiva-Duarte assessed study quality. João Locke Ferreira de Araújo, Diego Menezes, Julia Maria Saraiva-Duarte, and Renan Pedra de Souza drafted the manuscript. All authors revised and approved the final manuscript version.

DATA AVAILABILITY STATEMENT

All data is available upon request.

ORCID

João Locke Ferreira de Araújo https://orcid.org/0000-0001-9469-9673

Diego Menezes https://orcid.org/0000-0002-9603-455X

Julia Maria Saraiva-Duarte https://orcid.org/0000-0001-7742-3210

Luciana de Lima Ferreira https://orcid.org/0000-0002-4507-8261

Renato Santana de Aguiar https://orcid.org/0000-0001-5180-3717

Renan Pedra de Souza https://orcid.org/0000-0002-9479-4432

REFERENCES

1. Pifarré I Arolas H, Acosta E, López-Casasnovas G, et al. Years of life lost to COVID-19 in 81 countries. *Sci Rep*, 2021;11(1):3504. https://doi.org/10.1038/s41598-021-83040-3
2. Izcovich A, Ragusa MA, Tortosa F, et al. Prognostic factors for severity and mortality in patients infected with COVID-19: a systematic review. *PLoS ONE*. 2020;15(11). https://doi.org/10.1371/journal.pone.0241955
3. Sze S, Pan D, Nevill CR, et al. Ethnicity and clinical outcomes in COVID-19: a systematic review and meta-analysis. *Clin Med*. 2020;29-30. https://doi.org/10.1016/j.eclinm.2020.100630
4. Javanmardi F, Keshavarzi A, Akbari A, Emami A, Pirbonyeh N. Prevalence of underlying diseases in died cases of COVID-19: a systematic review and meta-analysis. *PLoS ONE*. 2020;15(10). https://doi.org/10.1371/journal.pone.0241265
5. Viner RM, Mytton OT, Bonell C, et al. Susceptibility to SARS-CoV-2 infection among children and adolescents compared with adults: a systematic review and meta-analysis. *JAMA Pediatrics*. 2021;175(2):143-156. https://doi.org/10.1001/jamapediatrics.2020.4573
6. Lakbar I, Luque-Paz D, Mege JL, Einav S, Leone M. COVID-19 gender susceptibility and outcomes: a systematic review. *PLoS ONE*. 2020;15(11). https://doi.org/10.1371/journal.pone.0241827
7. Macias Gil R, Marcelin JR, Zuniga-Blanco B, Marquez C, Mathew T, Piggott DA. COVID-19 pandemic: disparate health impact on the hispanic/latinx population in the United States. *J Infect Dis*. 2020;222(10):1592-1595. https://doi.org/10.1093/infdis/jiaa474
8. Lu R, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*. 2020;395(10224):565-574. https://doi.org/10.1016/S0140-6736(20)30251-8
9. Candido DS, Claro IM, de Jesus JG, et al. Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science*. 2020;369(6508). eabd2161. https://doi.org/10.1126/science.abd2161
10. Li Q, Wu J, Nie J, et al. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell*. 2020;182(5):1284-1294.e9. https://doi.org/10.1016/j.cell.2020.07.012
11. Davies NG, Jarvis CI, van Zandvoort K, et al. Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature*. 2021;593(7858). https://doi.org/10.1038/s41586-021-03426-1
12. Muus C, Luecken MD, Eraslan G, et al. Single-cell meta-analysis of SARS-CoV-2 entry genes across tissues and demographics. *Nat Med*. 2021;27(3):546-559. https://doi.org/10.1038/s41591-020-01227-z
13. Chan KK, Dorosky D, Sharma P, et al. Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science*. 2020;369(6508). https://doi.org/10.1126/SCIENCE.ABC0870
14. Buszko M, Nita-Lazar A, Park J-H, et al. Lessons learned: new insights on the role of cytokines in COVID-19. *Nat Immunol*. Published online 2021. https://doi.org/10.1038/s41590-021-00901-9
15. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. https://doi.org/10.1371/journal.pmed.1000097
16. Sohani ZN, Meyre D, de Souza RJ, Joseph PG, Gandhi MDB. Quality of genetic association studies (Q-Genie). Published online 2015:1-4.
17. Sohani ZN, Sarma S, Alyass A, et al. Empirical evaluation of the Q-Genie tool: a protocol for assessment of effectiveness. *BMJ Open*. 2016;6(6):1-7. https://doi.org/10.1136/bmjopen-2015-010403
18. Lorente L, Martín MM, Franco A, et al. HLA genetic polymorphisms and prognosis of patients with COVID-19. *Med Intensiva*. 2020;(xx). https://doi.org/10.1016/j.medin.2020.08.004
19. Littera R, Campagna M, Deidda S, et al. Human leukocyte antigen complex and other immunogenetic and clinical factors influence susceptibility or protection to SARS-CoV-2 infection and severity of the disease course. The Sardinian experience. *Front Immunol*. 2020;11:605688. https://doi.org/10.3389/fimmu.2020.605688
20. Gómez J, Albaiceta GM, García-Clemente M, et al. Angiotensin-converting enzymes (ACE, ACE2) gene variants and COVID-19

outcome. *Gene*. 2020;762:145102. https://doi.org/10.1016/j.gene.2020.145102
21. Wang F, Huang S, Gao R, et al. Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. *Cell Discov* 2020;6(1). https://doi.org/10.1038/s41421-020-00231-4
22. Rosenbaum JT, Hamilton H, Weisman MH, Reveille JD, Winthrop KL, Choi D. The effect of HLA-B27 on susceptibility and severity of COVID-19. *J Rheumatol*. Published online 2020:jrheum.200939. https://doi.org/10.3899/jrheum.200939
23. Ellinghaus D, Degenhardt F, Bujanda L, et al. Genomewide association study of severe covid-19 with respiratory failure. *N. Engl J Med*. Published online 2020. https://doi.org/10.1056/NEJMoa2020283
24. Kuo C-L, Pilling LC, Atkins JL, et al. APOE e4 Genotype Predicts Severe COVID-19 in the UK Biobank Community Cohort. *J Gerontol A Biol Sci Med Sci*. 2020;75(11):2231-2232. https://doi.org/10.1093/gerona/glaa131
25. Zhang Y, Zhang Y, Qin L, et al. Interferon-induced Transmembrane protein 3 genetic variant rs12252-C associated with disease severity in coronavirus disease 2019. *J Infect Dis*. 2020;222(1):34-37. https://doi.org/10.1093/infdis/jiaa224
26. Zhang Q, Bastard P, Liu Z, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science*. 2020;370(6515). https://doi.org/10.1126/SCIENCE.ABD4570
27. 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 Prepr Serv health Sci. Published online November 9, 2020. https://doi.org/10.1101/2020.11.05.20226761
28. Wang W, Zhang W, Zhang J, He J, Zhu F. Distribution of HLA allele frequencies in 82 Chinese individuals with coronavirus disease-2019 (COVID-19). *HLA*. 2020;96(2). https://doi.org/10.1111/tan.13941
29. Povysil G, Butler-Laporte G, Shang N, et al. Rare loss-of-function variants in type I IFN immunity genes are not associated with severe COVID-19. *J Clin Invest*. 131(4). https://doi.org/10.1172/jci147834
30. Novelli A, Andreani M, Biancolella M, et al. HLA allele frequencies and susceptibility to COVID-19 in a group of 99 Italian patients. *HLA*. 2020;96(5):610-614. https://doi.org/10.1111/tan.14047
31. Torre-Fuentes L, Matías-Guiu J, Hernández-Lorenzo L, et al. ACE2, TMPRSS2, and Furin variants and SARS-CoV-2 infection in Madrid, Spain. *J Med Virol*. 2020;(June):1-7. https://doi.org/10.1002/jmv.26319
32. Fan Q, Zhang W, Li B, Li DJ, Zhang J, Zhao F. Association between ABO blood group system and COVID-19 susceptibility in Wuhan. *Front Cell Infect Microbiol*. 2020;10:1-7. https://doi.org/10.3389/fcimb.2020.00404
33. Zhao J, Yang Y, Huang H, et al. Relationship between the ABO blood group and the COVID-19 susceptibility. *Clin Infect Dis*. 73(2):1-18. https://doi.org/10.1093/cid/ciaa1150
34. Izurieta HS, Graham DJ, Jiao Y, et al. Natural history of coronavirus disease 2019: risk factors for hospitalizations and deaths among >26 million US medicare beneficiaries. J Infect Dis. Published online December 2020. https://doi.org/10.1093/infdis/jiaa767
35. Ghamrawi R, Gunaratne M. COVID-19 and sex differences. *Mayo Clin Proc*. 2020;95(10):2189-2203. https://doi.org/10.1016/j.mayocp.2020.07.024
36. Maltezou HC, Raftopoulos V, Vorou R, et al. Association between upper respiratory tract viral load, comorbidities, disease severity, and outcome of patients with SARS-CoV-2 infection. J Infect Dis. Published online January 2021. https://doi.org/10.1093/infdis/jiaa804
37. Chen Y, Yang D, Cheng B, et al. Clinical characteristics and outcomes of patients with diabetes and COVID-19 in association with glucose-lowering medication. *Diabetes Care*. 2020;43(7):1399-1407. https://doi.org/10.2337/dc20-0660
38. Xu T, Chen C, Zhu Z, et al. Clinical features and dynamics of viral load in imported and non-imported patients with COVID-19. *Int J Infect Dis IJID*: official publication of the International Society for Infectious Diseases. 2020;94:68-71. https://doi.org/10.1016/j.ijid.2020.03.022
39. Huang Y, Paxton WA, Wolinsky SM, et al. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. *Nat Med*. 1996;2(11). https://doi.org/10.1038/nm1196-1240
40. Cheng Z, Zhou J, To KK-W, 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(8):1214-1221. https://doi.org/10.1093/infdis/jiv246
41. Wellington D, Laurenson-Schafer H, Abdel-Haq A, Dong T. IFITM3: how genetics influence influenza infection demographically. *Biomed J*. 2019;42(1):19-26. https://doi.org/10.1016/j.bj.2019.01.004
42. Lee N, Cao B, Ke C, et al. IFITM3, TLR3, and CD55 gene SNPs and cumulative genetic risks for severe outcomes in Chinese patients with H7N9/H1N1pdm09 influenza. *J Infect Dis*. 2017;216(1):97-104. https://doi.org/10.1093/infdis/jix235
43. Shokri P, Golmohammadi S, Noori M, et al. The relationship between blood groups and risk of infection with SARS-CoV-2 or development of severe outcomes: a review. *Rev Med Virol*. n/a(n/a). https://doi.org/10.1002/rmv.2247
44. Povysil G, Butler-Laporte G, Gharavi AG, Richards JB, Goldstein DB, Kiryluk K. Association of rare predicted loss-of-function variants of influenza-related type I IFN genes with critical COVID-19 pneumonia. Reply. *J Clin Invest*. Published online June 22, 2021. https://doi.org/10.1172/jci152475
45. Nakaoka H, Inoue I. Meta-analysis of genetic association studies: methodologies, between-study heterogeneity and winner's curse. *J Hum Genet*. 2009;54(11):615-623. https://doi.org/10.1038/jhg.2009.95
46. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *Eur J Hum Genet*. 2020;28(6). https://doi.org/10.1038/s41431-020-0636-6
47. The COVID-19 Host Genetics Initiative . Mapping the Human Genetic Architecture of COVID-19 by Worldwide Meta-Analysis. *Nature*. Published online 2021. https://doi.org/10.1038/s41586-021-03767-x
48. Shelton JF, Shastri AJ, Ye C, et al. Trans-ancestry analysis reveals genetic and nongenetic associations with COVID-19 susceptibility and severity. *Nat Genet*. 2021;53(6). https://doi.org/10.1038/s41588-021-00854-7
49. Kosmicki JA, Horowitz JE, Banerjee N, et al. Pan-ancestry exome-wide association analyses of COVID-19 outcomes in 586,157 individuals. *Am J Hum Genet*. 108(7). https://doi.org/10.1016/j.ajhg.2021.05.017
50. Naslavsky MS, Vidigal M, Matos LDRB, et al. Extreme phenotypes approach to investigate host genetics and COVID-19 outcomes. *Genet Mol Biol*. 2021;44(1). https://doi.org/10.1590/1678-4685-GMB-2020-0302
51. Little J, Higgins JPT, Ioannidis JPA, et al. STrengthening the REporting of genetic association studies (STREGA)- an extension of the STROBE statement. *Genet Epidemiol*. 2009;33(7):581-598. https://doi.org/10.1002/gepi.20410
52. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature*. 2016;538(7624). https://doi.org/10.1038/538161a
53. Novelli A, Biancolella M, Borgiani P, et al. Analysis of ACE2 genetic variants in 131 Italian SARS-CoV-2-positive patients. *Hum Genomics*. 2020;14(1). https://doi.org/10.1186/s40246-020-00279-z
54. Amodio E, Pipitone RM, Grimaudo S, et al. SARS-CoV-2 viral load, IFNλ polymorphisms and the course of COVID-19: an observational study. *J Clin Med*. 2020;9(10). https://doi.org/10.3390/jcm9103315
55. Benetti E, Tita R, Spiga O, et al. ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. *Eur J Hum Genet*. 2020;28(11). https://doi.org/10.1038/s41431-020-0691-z

56. Saleh A, Sultan A, Elashry Ma, et al. Association of TNF-α G-308 a Promoter Polymorphism with the Course and Outcome of COVID-19 Patients. *Immunol Invest*. Published online November 23, 2020. https://doi.org/10.1080/08820139.2020.1851709

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Ferreira de Araújo JL, Menezes D, Saraiva-Duarte JM, de Lima Ferreira L, Santana de Aguiar R, Pedra de Souza R. Systematic review of host genetic association with Covid-19 prognosis and susceptibility: what have we learned in 2020? *Rev Med Virol*. 2022;32(2):e2283. https://doi.org/10.1002/rmv.2283