UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE GENÉTICA, ECOLOGIA E EVOLUÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA

JOÃO LOCKE FERREIRA DE ARAÚJO

ASSOCIAÇÃO DE BIOMARCADORES GENÉTICOS DO HOSPEDEIRO COM O PROGNÓSTICO E A SUSCEPTIBILIDADE DA COVID-19

BELO HORIZONTE
2023

João Locke Ferreira de Araújo

ASSOCIAÇÃO DE BIOMARCADORES GENÉTICOS DO HOSPEDEIRO COM O PROGNÓSTICO E A SUSCEPTIBILIDADE DA COVID-19

Tese submetida ao programa de Pós-graduação em Genética da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do título de Doutor em Genética.

Orientador: Dr. Renan Pedra de Souza
Coorientador: Dr. Renato Santana de Aguiar

Área de concentração: Genética evolutiva e de populações.

BELO HORIZONTE
2023

043 Araújo, João Locke Ferreira de.
Associação de biomarcadores genéticos do hospedeiro com o prognóstico e a susceptibilidade da COVID-19 [manuscrito] / João Locke Ferreira de Araújo. – 2023.
107 f. : il. ; 29,5 cm.

Orientador: Dr. Renan Pedra de Souza. Coorientador: Dr. Renato Santana de Aguiar.
Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Genética. 2. Infecções por Coronavirus. 3. Biomarcadores. 4. Metanálise. I. Souza, Renan Pedra de. II. Aguiar, Renato Santana de. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 575

Ficha catalográfica elaborada pela Bibliotecária Fabiane C M Reis. CRB/6 - 2680

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Genética

ATA DE DEFESA DE TESE

ATA DA DEFESA DE TESE	**170/2023** **entrada**
João Locke Ferreira de Araújo	**1º/2018** **CPF: 099.811.766-80**

Às quatorze horas do dia **31 de janeiro de 2023**, reuniu-se, a Comissão Examinadora de Tese, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: **"Associação de biomarcadores genéticos do hospedeiro com o prognóstico e a susceptibilidade da COVID-19"**, requisito para obtenção do grau de Doutor em **Genética.** Abrindo a sessão, o Presidente da Comissão, **Renan Pedra de Souza**, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

Prof./Pesq.	**Instituição**	**CPF**	**Indicação**
Renan Pedra de Souza	UFMG	064.488.066-01	APROVADO
Ana Lúcia Brunialti Godard	UFMG	107.961.538-50	APROVADO
Elaine Virgínia Martins de Souza Figueiredo	UFAL	041.392.544-75	APROVADO
Anna Carolina Toledo da Cunha Pereira	Universidade Federal do Delta do Parnaíba (UFDPar)	003.094.646-84	APROVADO
Rodrigo Araújo Lima Rodrigues	UFMG	103.110.986-22	APROVADO

Pelas indicações, o candidato foi considerado: APROVADO

O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 31 de janeiro de 2023.

Renan Pedra de Souza

Ana Lúcia Brunialti Godard

Elaine Virgínia Martins de Souza Figueiredo

Anna Carolina Toledo da Cunha Pereira

Rodrigo Araújo Lima Rodrigues

Documento assinado eletronicamente por **Elaine Virgínia Martins de Souza Figueiredo**, **Usuária Externa**, em 02/02/2023, às 10:39, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Rodrigo Araújo Lima Rodrigues**, **Professor do Magistério Superior**, em 02/02/2023, às 10:54, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Anna Carolina Toledo da Cunha Pereira**, **Usuária Externa**, em 02/02/2023, às 14:50, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Renan Pedra de Souza**, **Professor do Magistério Superior**, em 02/02/2023, às 18:23, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Ana Lucia Brunialti Godard**, **Professora do Magistério Superior**, em 06/02/2023, às 09:40, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2057460** e o código CRC **E63A5755**.

Referência: Processo nº 23072.204706/2023-03 SEI nº 2057460

UNIVERSIDADE FEDERAL DE MINAS GERAIS
[NOME DA UNIDADE]
COLEGIADO DO CURSO DE GRADUAÇÃO / PÓS-GRADUAÇÃO EM [XXXX]

FOLHA DE APROVAÇÃO

"Associação de biomarcadores genéticos do hospedeiro com o prognóstico e a susceptibilidade da COVID-19"

João Locke Ferreira de Araújo

Tese aprovada pela banca examinadora constituída pelos Professores:

Renan Pedra de Souza
UFMG

Ana Lúcia Brunialti Godard
UFMG

Elaine Virgínia Martins de Souza Figueiredo
UFAL

Anna Carolina Toledo da Cunha Pereira

Universidade Federal do Delta do Parnaíba (UFDPar)

Rodrigo Araújo Lima Rodrigues

UFMG

Belo Horizonte, 31 de janeiro de 2023.

Documento assinado eletronicamente por **Elaine Virgínia Martins de Souza Figueiredo**, **Usuária Externa**, em 02/02/2023, às 10:40, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Rodrigo Araújo Lima Rodrigues**, **Professor do Magistério Superior**, em 02/02/2023, às 10:55, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Anna Carolina Toledo da Cunha Pereira**, **Usuário Externo**, em 02/02/2023, às 14:50, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Renan Pedra de Souza**, **Professor do Magistério Superior**, em 02/02/2023, às 18:23, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Ana Lucia Brunialti Godard**, **Professora do Magistério Superior**, em 06/02/2023, às 09:40, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2057496** e o código CRC **E2AA603C**.

Referência: Processo nº 23072.204706/2023-03 SEI nº 2057496

Aos 698.018 brasileiros mortos pela COVID-19, e suas famílias.

Última atualização em 16 de fevereiro de 2023 *(https://covid.saude.gov.br/).*

AGRADECIMENTOS

Agradeço aos meus pais, Enirson e Geralda, e ao meu irmão Francisco, por toda a confiança e apoio que sempre pude contar, e por nunca duvidarem de minha capacidade.

Aos meus colegas de laboratório Diego, Hugo, Fernanda Souza, Isabela, Luiza, Victor, Luciene, Rafa, Daniel, Paula, Julia, Wallyson, Rillery, Fran, Fernanda Marin, João Victor, Aryel e Ana, por terem tornado cada minuto passado no LBI o mais agradável possível, o que espero que se mantenha até o fim desta trajetória. Hugo em especial, por toda a parceria em experimentos e trocas de conhecimentos sobre word e excel.

Ao meu orientador Professor Renan P. de Souza, por toda a paciência e compreensão, além da enorme dedicação e preocupação para contribuir para a minha formação como profissional (como ele mesmo diz, preocupação com a formação de recursos humanos). Por ter abraçado minhas ideias, e encorpado ou reforçado com outras. Sei que ainda irei aprender muito com ele no tempo em que me resta no caminho até ao título.

À Lourdes, pelo carinho incondicional e por toda a orientação e apoio que me deu nos momentos mais difíceis.

Aos meus amigos da graduação, da pós e da vida, Pablo e Fernanda, por estarem sempre presentes em cada momento.

Ao Bruno, Renata e Marina, pelo apoio emocional, amizade e descontração. Bruno e Renata em especial por terem me ajudado a abandonar o sedentarismo que me acompanhou por boa parte da pós-graduação.

Ao Fabio pela amizade, ombro amigo, carinho, parceria e aulas de inglês. Você é um cara excepcional.

Ao Rafael, Isabela e Alice. Rafael pela amizade sincera de tantos anos, e sua família que hoje é minha também.

A Priscila e ao João Pedro. Priscila em especial pelos quase 16 anos de amizade.

A Laurinha e ao Fê, casal de pesquisadores incríveis que me inspiram.

Ao grupo ‘Biologando’ pela inspiração e por possibilitar que eu jamais perca o amor pela profissão.

Ao meu ex coorientador de doutorado, Maicon Albuquerque, por ter abraçado meu projeto inicial e possibilitado que eu fosse aprovado na seleção de doutorado. Projeto esse que não desisto de colaborar e ajudar a colocá-lo em prática em breve. Meus olhos brilharam com a genética do esporte quando vi a enorme possibilidade de se trabalhar biologia evolutiva e registro aqui meu desejo e disponibilidade em colaborar.

Ao meu atual coorientador Renato Santana, por toda a ajuda e orientação prestada durante o tempo que venho trabalhado no LBI na pandemia da COVID-19, onde me vi obrigado a

mudar de projeto e trabalhar minha base teórica do zero. Aprendi em tempo recorde a trabalhar com vírus, assim como meus colegas, graças ao seu enorme background e o prazer em ensinar.

Aos amigos da panela do IEMG, Leandro, Wellerson e Christiano por jamais deixarem me esquecer de onde eu vim e por estarem comigo para onde eu irei.

Ao Eduardo, Lara e Davi, cuja simples existência me permite jamais esquecer de me amar mais, pois sou amado demais.

Ao meu mestre de Taekwondo, Carlos Franco, e sua esposa Conceição, pela minha segunda família e por toda confiança depositada em mim.

À toda família taekwondo, em especial a família que construí na arbitragem por todos os momentos excepcionais que vivemos no fomento do esporte que tanto amamos.

À minha noiva e companheira Gracielle Braga, por ser meu porto seguro e meu maior exemplo de amor pelo trabalho e dedicação por tudo o que se presta a fazer e a defender. Obrigado, meu amor, por ser meu maior presente.

Ao Bernardo e a Sofia, por terem me dado o gosto da paternidade, mesmo não biológica, e o sentimento de ser amado por um filho.

À Heveline, por todo o carinho com o amor da minha vida, e comigo também.

Agradeço a toda a comunidade cientifica e as vacinas, sem as quais a dedicatória desta tese teria abrangido muito mais pessoas.

A todos aos meus amigos e familiares que não citei aqui, mas que merecem um agradecimento especial simplesmente por fazerem parte da minha vida.

“In the long history of humankind (and animal kind, too) those who learned to collaborate and improvise most effectively have prevailed.”

Charles Darwin

“The sweetest and most inoffensive path of life leads through the avenues of science and learning; and whoever can either remove any obstructions in this way, or open up any new prospect, ought so far to be esteemed a benefactor to mankind.”

David Hume

“The fact that a opinion has been widely held is no evidence whatever that it is not utterly absurd; indeed in view of the silliness of the majority of mankind, a widespread belief is more likely to be foolish than sensible”

Bertrand Russell

“Faz do jeito certo primeiro! Faz errado só se sobrar tempo! ”

Tião Piquete (Sebastião Ferreira, meu avô)

RESUMO

A doença do coronavírus 2019 (COVID-19) é causada por um vírus da família Coronaviridae, conhecido como o Coronavírus da Síndrome Respiratória Aguda Grave 2 (SARS-CoV-2). A manifestação clínica é bastante heterogênea, sendo que os infectados podem variar desde assintomáticos a sintomas respiratórios graves, podendo vir à óbito. Neste trabalho exploramos biomarcadores genéticos que podem ser associados a variabilidade nos desfechos clínicos da COVID-19, assim como a susceptibilidade à infecção pelo SARS-CoV-2. No primeiro capítulo, realizamos uma revisão sistemática envolvendo estudos de associação genética e o prognóstico e susceptibilidade da COVID-19 publicados durante o ano de 2020. Foram inclusos 20 estudos, sendo 14 de susceptibilidade e 11 de prognóstico, com 5 abordando ambos os desfechos. No capítulo 2, realizamos uma revisão sistemática seguida de metanálise para explorar a contribuição dos genes *IFITM3*, *TNF-alpha*, *ACE1* e *FURIN* no prognóstico da COVID-19. Encontramos 5 estudos para o *IFITM3*, 3 para o *TNF-alpha*, 17 para o *ACE1* e 2 para o *FURIN*. A metanálise foi possível para o polimorfismo rs12252 no *IFITM3*, polimorfismo rs4646994 no *ACE1 e* polimorfismo rs1800629 no gene *TNF-alpha.* Nenhum efeito foi observado para o polimorfismo rs12252 no *IFITM3 na severidade* e para o polimorfismo rs1800629 no *TNF-alpha* com a chance de óbito*.* Contudo, um efeito alélico de risco para o alelo deletério (alelo D) do polimorfismo rs4646994 no *ACE1* foi observado (OR: 1,45; IC 95%: 1,26 – 1,66), além de um efeito genotípico do genótipo D/D para a gravidade (OR: 1,49, 95% CI:1,22 – 1,83), com o genótipo homozigoto para a inserção (I/I) promovendo um efeito protetor (OR: 0,57, 95% CI: 0,45 – 0,74). No capítulo 3, investigamos a influência da idade, sexo e dos níveis de expressão dos genes *ACE1*, *ACE2* e *TMPRSS2*, além do polimorfismo rs4646994 no aumento da chance do uso de ventilação mecânica e na chance do óbito por COVID-19. Observamos associação da idade com a chance de internação (p<0,001) na coorte do Rio de Janeiro. Nenhuma alteração na expressão dos genes *ACE1*, *ACE2* e *TMPRSS2* foi associada a gravidade ou óbito. O polimorfismo rs4646994 no gene *ACE1* também não apresentou influência para nenhum dos resultados. Tambem não se observou associação na coorte do Rio de Janeiro para hospitalização, embora tenha-se observado efeito em uma metanálise combinada com a literatura: alelo D dominante (OR:1,39; 95%CI: 1,12- 1,72) e alelo I dominante (OR:0,76; 95%CI: 0,61-0,95). No capítulo 4 tentamos replicar o lócus 3p21.31 reportado como associado a COVID-19 grave com polimorfismos nos genes *CXCR6* e *LZTFL1.* Nenhuma associação foi observada para os polimorfismos *CXCR6* (rs2234358) e *LZTFL1* (rs10490770) tanto para o suporte respiratório quanto a chance de

óbito. Contudo, uma associação foi observada para o polimorfismo *CXCR6* (rs2234355) para a chance de óbito (p=0,022) em um modelo de codominância, e o genótipo AG apresentou uma diminuição da chance de óbito (OR: 0,09; 95% CI: 0,011-0,576). No capítulo 5 reportamos em uma carta os resultados encontrados em uma análise de associação envolvendo o polimorfismo rs12252 no gene *IFITM3* na dificuldade respiratória e na chance de óbito. Nenhuma associação foi observada para nenhum dos desfechos. Estes resultados reforçam a importância da replicação dos estudos de associação genética na COVID-19. O uso de metanálises para aumentar a confiabilidade dos resultados é de extrema importância, assim como o cuidado com a qualidade dos estudos originais publicados. É fundamental que mais biomarcadores no prognóstico sejam explorados para que possamos chegar a cada vez mais próximos de compreender a heterogeneidade dos sintomas na COVID-19. Considerando a sintomatologia de doenças infeciosas multifatorial, compreender a dinâmica de biomarcadores nos desfechos clínicos é fundamental para intervenções seguras da doença, e estudos envolvendo a influência de múltiplos marcadores genéticos torna-se extremamente importante nesse contexto.

Palavras chaves: genética de associação; polimorfismos; SARS-CoV-2; COVID-19; prognóstico, suscetibilidade

ABSTRACT

Coronavirus disease 2019 (COVID-19) is caused by a virus in the family Coronaviridae, known as the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). The clinical manifestation is quite heterogeneous, and those infected can range from asymptomatic to severe respiratory symptoms and may even die. In this work, we explore genetic biomarkers that may be associated with variability in clinical outcomes of COVID-19, as well as susceptibility to SARS-CoV-2 infection. In the first chapter, we performed a systematic review involving genetic association studies and the prognosis and susceptibility of COVID-19 published during the year 2020. 20 studies were included, 14 on susceptibility and 11 on prognosis, with 5 addressing both outcomes. In Chapter 2, we performed a systematic review followed by a meta-analysis to explore the contribution of *IFITM3*, *TNF-alpha*, *ACE1* and FURIN genes to the COVID-19 prognosis. We found 5 studies for *IFITM3*, 3 for *TNF-alpha*, 17 for *ACE1* and 2 for *FURIN*. Meta-analysis was possible for the rs12252 polymorphism in *IFITM3*, rs4646994 polymorphism in *ACE1* and rs1800629 polymorphism in the *TNF-alpha* gene. No effect was observed for the rs12252 polymorphism in *IFITM3* on severity and for the rs1800629 polymorphism in *TNF-alpha* on mortality. However, an allelic risk effect for the deleterious allele (D allele) of the rs4646994 polymorphism in *ACE1* was observed (OR: 1.45; 95% CI: 1.26 – 1.66), in addition to a genotypic effect of genotype D /D for severity (OR: 1.49, 95% CI:1.22 – 1.83), with the homozygous genotype for insertion (I/I) promoting a protective effect (OR: 0.57, 95% CI: 0.45 – 0.74). In chapter 3, we investigated the influence of age, sex and the *ACE1*, *ACE2* and *TMPRSS2* genes expression levels, in addition to the rs4646994 polymorphism, on the increased chance of using mechanical ventilation and the chance of death from COVID-19. We observed an association between age and the chance of hospitalization ($p<0.001$) in the Rio de Janeiro cohort. No change in *ACE1*, *ACE2* and *TMPRSS2* gene expression was associated with severity or mortality. The rs4646994 polymorphism in the *ACE1* gene also did not influence any of the results. No association was observed in the Rio de Janeiro cohort for hospitalization either, although an effect was observed in a meta-analysis combined with the literature: dominant D allele (OR:1.39; 95%CI: 1.12-1.72) and dominant I allele (OR:0.76; 95%CI: 0.61-0.95). In chapter 4 we attempted to replicate the 3p21.31 locus reported to be associated with severe COVID-19 with polymorphisms in the *CXCR6* and *LZTFL1* genes. No association was observed for the *CXCR6* (rs2234358) and *LZTFL1* (rs10490770) polymorphisms for either respiratory support or mortality. However, an association was observed for the *CXCR6* polymorphism (rs2234355) for mortality ($p=0.022$) in a codominance model, and the AG genotype showed

a decreased chance of death (OR: 0.09; 95% CI: 0.011 -0.576). In chapter 5, we report in a letter the results found in an association analysis involving the rs12252 polymorphism in the *IFITM3* gene in respiratory distress and chance of death. No association was observed for any of the outcomes. These results reinforce the importance of replicating genetic association studies in COVID-19. The use of meta-analyses to increase the reliability of results is extremely important, as is care for the quality of published original studies. It is essential that more prognostic biomarkers are explored so that we can come ever closer to understanding the heterogeneity of symptoms in COVID-19. Considering the multifactorial symptomatology of infectious diseases, understanding the dynamics of biomarkers in clinical outcomes is essential for safe disease interventions, and studies involving the influence of multiple genetic markers become extremely important in this context.

Keywords: genetics association; polymorphisms; SARS-CoV-2; COVID-19; prognosis, susceptibility

LISTA DE FIGURAS

LISTA DE ABREVIAÇÕES

ACE1 - angiotensin I converting enzyme

ACE2 - angiotensin converting enzyme II

CCR9 - C-C motif chemokine receptor 9

CI - Confidence interval

COVID-19 - coronavirus disease 2019

CTSL - cathepsin L

CXCR6 - C-X-C motif chemokine receptor 6

DPP4 - dipeptidyl peptidase 4

FURIN - paired basic amino acid cleaving enzyme

FYCO1 – FYVE and coiled-coil domain autophagy adaptor 1

GM-CSF – Granulocyte´-macrophage colony-stimulating factor

GWAS - Genome Wide Association Studies

HIV - Human immunodeficiency virus

HLA - Human Leukocyte Antige

ICB - Instituto de Ciências Biológicas

IFITM3 - interferon induced trasmembrane protein 3

IFN - interferon

IFNL3 - interferon lambda 3

IFNL4 - interferon lambda 4

IL-6 - interleukin 6

LBI - Laboratório de Biologia Integrativa

LZTFL1 - leucine zipper transcription factor like 1

MERS - Middle East Respiratory Syndrome

NGS - New Generation Sequencing

OMS - Organização mundial de saúde

OR - Odds Ratio

PCSK5 - proprotein convertase subtilisin/kexin type 5

PCSK7 - proprotein convertase subtilisin/kexin type 7

PRISMA - Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Q-Genie - Quality of genetic association studies

RBD - receptor-binding domain

SARS-CoV-2 - Severe Acute Respiratory Syndrome Coronavirus 2

SLC6A20 - Solute Carrier Family 6 member 20

SNP - Single nucleotide polymorphism

STREGA - Strengthening the REporting of Genetic Association studies

TGF-BETA - Transforming growth factor beta

TMPRSS2 - transmembrane serine protease 2

TNF-α - Tumor necrosis factor alpha

UFMG - Universidade Federal de Minas Gerais

VOC - Variant of concern

VOI - Variant of intersting

XCR1 - X-C motif chemokine receptor 1

Sumário

INTRODUÇÃO

COVID-19: Aspectos gerais

Pandemia

A doença do coronavírus 2019 (COVID-19) é causada por um vírus da família Coronaviridae, conhecido como o Coronavírus da Síndrome Respiratória Aguda Grave 2 (SARS-CoV-2)[1]. Coronaviridae é uma família de vírus bastante diversa [2,3]. Ela é formada por vírus com genoma baseado em RNA de fita simples e com sentido positivo[4,5]. Eles apresentam um desafio na saúde pública, já que a infecção por esses vírus resulta principalmente em doenças respiratórias e entéricas[6,7]. Nas duas últimas décadas, houve dois outros surtos causados por coronavírus, o SARS-CoV (2002) e o MERS (2012), além da atual pandemia por SARS-CoV-2[8]. O primeiro registro da COVID-19 ocorreu em novembro de 2019, na cidade de Wuhan, na China[9,10]. A evidência de transmissão entre humanos tornou-se oficializada em janeiro de 2020, após uma visita realizada pela OMS à cidade de Wuhan[11]. Desde o primeiro surto reconhecido em fevereiro de 2020, a doença vem se espalhando pelo mundo, sendo decretada crise sanitária global e pandemia pela Organização Mundial de Saúde (OMS), em 11 de março de 2020. De acordo com a OMS, em janeiro de 2023 registrou-se 7,507 milhões de mortes em todo o mundo. No Brasil, o número de óbitos ultrapassou 600 mil no início de outubro de 2021, e em janeiro de 2023 contabiliza 694 mil óbitos. Assim como outros coronavírus recentes, ao infectar células epiteliais brônquicas e células do trato respiratório em humanos, a infecção pode evoluir para casos graves e óbito[6,12,13].

Atualmente existem várias vacinas disponíveis no mercado. Apesar da dinâmica de vacinação variar entre os países, no final de 2022 cerca de 70% da população mundial já se encontra vacinada com pelo menos 1 dose e 64,6% totalmente vacinadas com 2 doses. No Brasil, 81,7% da população já se encontra vacinada com duas doses em dezembro de 2022. Vacinas utilizando partículas virais inativadas, assim como abordagens com adenovírus expressando proteínas virais e até mesmo vacinas envolvendo tecnologia de RNAm já foram e estão sendo utilizadas, e observamos resultados satisfatórios à medida que a vacinação avançou. A vacinação em massa já foi associada com menores taxas de mortalidade e de internações, além de contribuir para diminuir a transmissão do virus[14–16].

Devida a alta taxa de mutação de genomas virais, o aparecimento de variantes dos SARS-CoV-2 tornou-se comum à medida que a transmissão e o espalhamento não são

controlados[16]. A proteína do envelope viral, denominada Spike, apresenta uma variabilidade considerável se comparada a outras proteínas. Isso ocorre porque esta proteína sofre uma pressão seletiva maior do ambiente por estar diretamente relacionada a entrada do vírus na célula, fazendo com que seja observada grande variabilidade no gene S (Spike) entre as variantes já consolidadas no ambiente. Estas variações podem levar a otimização do processo de entrada do virus, além do escape a resposta imunológica devida ao não reconhecimento por anticorpos neutralizantes[17,18].

O Brasil ficou marcado por três grandes ondas de transmissão, sendo que na segunda registrou-se o aparecimento das variantes de interesse (variants of intersting - VOI) e das variantes de preocupação (variants of concern - VOC). As VOI representam linhagens que possuem mutações estruturais que podem afetar diretamente a dinâmica da doença, sendo capaz de influenciar nas taxas de mortalidade, na gravidade dos sintomas e até mesmo atribuir menor resposta as vacinas já disponíveis[19]. As VOC, por sua vez, possuem estas mutações já associadas a um ou mais destes desfechos, já devidamente reportados em estudos científicos[19]. A variante Alfa (antiga B.1.1.7), descrita primeiramente na Inglaterra está associada à maior gravidade no desfecho clínico da doença [20]. As variantes Zeta (Rio de Janeiro) e Gama (Manaus) foram descritas no Brasil em março de 2021 e são originadas da linhagem VOI B.1.1.28, tendo a variante Zeta surgido primeiro[18,21,22]. Há evidências também de reinfecção por estas variantes[23,24]. O surgimento e a dispersão global das VOC foram considerados o principal fator que levou a consolidação da pandemia pelo mundo[25].

Variantes tais como as variantes Gama (Brasil) e Delta (Índia), hoje não mais circulantes no Brasil, foram relacionadas com a maior gravidade de sintomas assim como a transmissão do SARS-CoV-2[26]. De uma forma geral, as variantes classificadas como VOC estão associadas à gravidade dos sintomas na COVID-19, além de maior transmissibilidade e ao escape vacinal. Por conta disso, grupos de pesquisa pelo país tem investido na vigilância genômica das variantes do SARS-CoV-2 de modo que a compreensão da dinâmica populacional destas variantes auxilie em métodos de intervenção e combate a pandemia[27].

A variante delta começou a se espalhar pelo Brasil no segundo semestre de 2021 [28]. O primeiro caso reportado foi em maio de 2021 no estado do maranhão[29]. Contudo, não observamos um aumento do número de casos. Muito deve-se a alta taxa de cobertura vacinal em que a população se encontrava no momento, além da imunidade natural adquirida por infecções previas pela variante Gama, responsável pela segunda onda de infecções no País [30]. Uma terceira onda de infecções aconteceu em janeiro de 2022 devida

a introdução da variante Ômicron e subvariantes[31]. Ela se espalhou rapidamente pelo mundo sendo que nas primeiras 3 semanas já havia sido identificada em 87 países[32]. A variante Ômicron tem se destacado pelo surgimento de subvariantes, como por exemplo as subvariantes BA.1 e BA.2, sendo esta última mais transmissível devida a mutações na RBD da proteína Spike[33]. Algo que com certeza pode ter impulsionado a dispersão destas subvariantes é o escape vacinal promovido pelas novas mutações no genoma da Ômicron, onde elas têm apresentado resistência aos anticorpos produzidos após infecções pela por primeiras cepas da variante[34]. Até dezembro de 2022, a subvariante da Ômicron prevalente no Brasil é a BQ.1, possuindo grande escape imunológico[35].

Clínica

Abordagens imunoterapêuticas para o tratamento clínico já foram propostas. A terapia com anticorpos neutralizantes monoclonais é uma alternativa para doenças infecciosas, e abordagens semelhantes já são usadas para o SARS-CoV e adaptadas para o SARS-CoV-2[36]. Estudos recentes mostraram que houve redução do risco de hospitalização em pacientes diagnosticados com COVID-19 que foram tratados com anticorpos neutralizantes além da redução do risco de ser infectado pelo vírus[37,38]. O risco de hospitalização foi de até 4.8% menor em pacientes que receberam a terapia com anticorpos neutralizantes do que o grupo com o tratamento placebo, além de observada redução significativa da carga viral[37,38]. O uso de anticorpos monoclonais é bastante limitado no Brasil.

Os estudos disponíveis até o momento apontam que cerca de um terço das infecções pelo SARS-CoV-2 são assintomáticas, além de estudos longitudinais mostrarem que aproximadamente 75% de pacientes com perfil assintomático no momento do teste permanecem sem apresentar sintomas durante o curso da infecção [39]. Contudo, mesmo sem sintomas, estas pessoas continuam sendo potenciais transmissores da doença[39,40]. Essas evidências crescentes ressaltam a importância de medidas sanitárias gerais, além de um sistema de vigilância adequado para que a transmissão possa ser controlada[41–44]. Devida a dificuldade de se rastrear casos assintomáticos, o controle da pandemia torna-se comprometido[42].

Já os pacientes sintomáticos podem evoluir para sintomas mais graves e até mesmo vir a falecer [45]. A sintomatologia da doença clínica é bastante ampla, sendo febre, tosse e dispneia sintomas considerados os mais frequentes[46,47]. Mesmo assim, eles podem não estar presentes, dificultando a definição do caso. É comum a queixa de sintomas gastrointestinais e perda de paladar (disgeusia) ou olfato (anosmia) entre pacientes em

casos considerados leves[48,49]. Os sintomas de dispneia são bastante relatados entre os casos graves [50,51] (Figura 1).

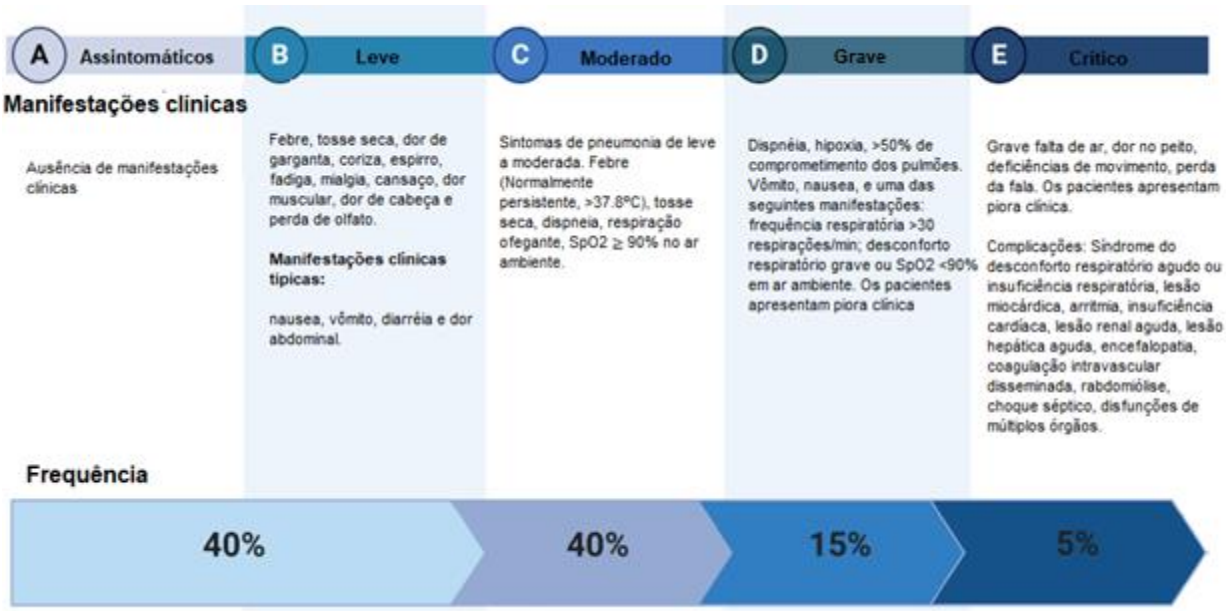


Figura 1 Adaptado de Galindo e Valencia, 2021. Curso clínico da COVID-19. Cerca de 5% das pessoas infectadas chegam ao estado crítico

O SARS-CoV-2 é transmitido através das vias aéreas superiores, e algumas evidências têm apontado para contaminação via fecal-oral[52,53]. A transmissão por partículas aerotransportadas demostrou enorme virulência e tem dominado a transmissão da COVID-19 [54–56]. Até o momento, as melhores medidas a serem tomadas para conter a transmissão é evitar a exposição ao vírus [57]. Desde o início da pandemia existiram recomendações sanitárias básicas, como lavar as mãos frequentemente, manter o isolamento social (evitar contato próximo com as pessoas, de preferência físico), utilização de máscara cobrindo adequadamente a boca e o nariz em ambientes compartilhados, sendo que o uso de máscara é considerado o meio mais eficaz de se evitar contaminação entre pessoas, cobrir tosses e espirros e higienizar superfícies comumente usadas diariamente[57,58].

Fisiopatologia da COVID-19

As primeiras etapas da infecção pelo SARS- CoV-2 envolvem uma ligação especifica da proteína spike (S) do vírus com receptores de entrada celular, sendo o mais comum o receptor ACE2 (enzima conversora de angiotensina 2)[59,60] (Figura 2).

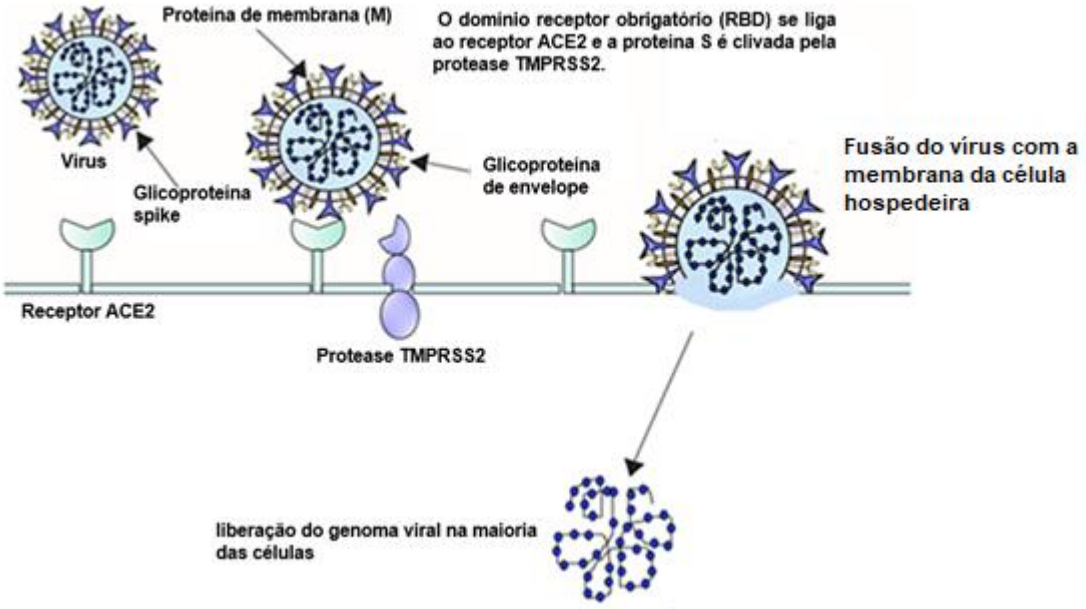


Figura 2 Adaptado de Mol. Pharmaceutics 2021 - Interação do receptor do hospedeiro com a proteína spike no SARS-CoV-2 e subsequente fusão do vírus com a membrana da célula hospedeira.

Compreender a dinâmica da interação vírus-célula hospedeira é importante para entender vários aspectos da doença, considerando que os receptores de entrada celular são fatores cruciais para determinar a afinidade do vírus com a célula e responsáveis por influenciar a gravidade das infecções por vírus específicos [3]. A alta taxa de mutações em que os vírus estão sujeitos pode interferir na especificidade ou afinidade da ligação do vírus a um receptor específico [4,10,61]. Doenças geralmente leves do trato respiratório causadas por vírus diferentes do SARS-CoV-2, mas que usam o mesmo receptor ACE2, podem ser explicadas com sua baixa interação ou afinidade pelo receptor[62]. Um exemplo é o vírus CoV-NL63, que causa transtornos respiratórios leves[62]. Dada a alta semelhança estrutural entre as proteínas S do SARS-CoV e SARS-CoV-2, o receptor ACE2 funciona de forma similar para ambos[60]. Contudo, evidências biofísicas e estruturais demostram que a afinidade da ligação do receptor ACE2 à proteína S do SARS-CoV-2 é 10-20 vezes maior do que a afinidade pela proteína S do SARS-CoV[8].

A invasão das células do hospedeiro pelo vírus não depende apenas da ligação da proteína S do vírus com o receptor, mas também requer uma clivagem da proteína por uma

protease[63]. Outra interação é de extrema importância para a infecção: a ligação entre o SARS-CoV-2 e a protease transmembranar serina 2, a TMPRSS2, já que ela cliva o domínio receptor obrigatório (RBD) da proteína Spike expondo o peptídeo de fusão, o que é crucial para a entrada do vírus na célula hospedeira (Figura 1) [63–65]. As serinas proteases são conhecidas por estarem envolvidas em inúmeros processos patológicos e biológicos[66,67]. Observou-se que o *TMPRSS2* é altamente expresso e em um maior número de células em comparação com o *ACE2*[63]. Isso sugere que o ACE2 poderia funcionar como um fator de limitação da taxa de entrada do vírus durante o estágio de infecção[40,68]. A TMPRSS2 é regulada por andrógenos, sendo bastante expressa no tecido gonadal e próstata[67]. Ela possui um papel significativo na regulação de canais de sódio, e é uma protease em potencial de ativação para a entrada celular em diversas infecções virais[69]. Além desta protease ser alvo para o desenvolvimento de terapias contra infecções causadas por SARS e MERS, também foi considerada como alvo terapêutico na COVID-19, no desenvolvimento de inibidores[70].

Mesmo que o *TMPRSS2* esteja consolidado como peça-chave na infecção por SARS-CoV-2, outras proteases também podem atuar na clivagem da proteína S, dado que o SARS-CoV-2 infecta células em ausência de TMPRSS2[71]. Portanto, proteases adicionais desempenham papeis na clivagem proteolítica das proteínas virais na entrada celular. Um exemplo é a catepsina L (CTSL), além de outras catepsinas, que podem atuar como substituta da TMPRSS2 na infecção[63]. A catepsina L é uma enzima lisossomal que participa de vários processos fisiológicos, incluindo apoptose, remodelação da matriz extracelular e processamento de antígenos[72–74]. Durante o processo de inflamação crônica ocorre a regulação da expressão da CTSL e esse processo possui envolvimento com a degradação da matriz extracelular, importante para a entrada do SARS-CoV-2 na célula do hospedeiro[73]. A catepsina L é um alvo terapêutico considerável em estudos recentes pela sua inibição poder interromper a infecção[72].

Outras proteases como FURIN, PCSK5 e PCSK7 são mais amplamente expressas do que TMPRSS2 em todos os tipos de células pulmonares[71]. A furina, por exemplo, é uma endoprotease celular que catalisa a ativação proteolítica de substratos proteicos em compartimentos de vias secretoras[73]. Na virologia, ela também possui um papel interessante, já que muitos vírus patogênicos, como os vírus da gripe aviária, HIV e sarampo expressam glicoproteínas de envelope que devem ser clivadas pela furina de complexo de Golgi para maturação da partícula viral [75–77]. Recentemente mostrou-se que o SARS-CoV-2 também possui um sítio de clivagem em potencial para furina[78]. Hoffmann e colaboradores

mostraram que o SARS-CoV-2 depende da pré clivagem mediada pela furina da proteína S no sítio S1/S2 para a posterior ativação pela TMPRSS2 em células pulmonares[78].

Um crescente corpo de dados clínicos sugere que uma tempestade de citocinas é associada à gravidade da doença e é um fator importante que pode levar ao óbito. Sabe-se que ela é observada em casos mais graves de inúmeras doenças infecciosas, incluindo SARS e MERS[79]. O SARS-CoV-2 induz uma resposta imunológica com produção de citocinas inflamatórias desencadeadas por interferon (IFN)[80,81]. Com o volume de citocinas elevado, ocorre uma migração de células de defesa para dentro do tecido, como macrófagos e neutrófilos, resultando na tempestade de citoninas[82,83]. O SARS-CoV-2 pode ativar rapidamente células produtoras de interferon gama (Th1) para secretar citocinas pró inflamatórias, como fatores de estimulação de colônias de granulócitos e macrófagos (GM-CSF) e interleucina 6 (IL-6)[81]. Monócitos inflamatórios CD14+CD16+ são ativados para produzir grandes quantidades de IL-6, fator de necrose tumoral-alpha (TNF- ***α***) e outras citocinas[79,80,83].

Biomarcadores na COVID-19

Um biomarcador é um parâmetro que é utilizado como um indicador para um processo biológico normal ou patogênico ou como uma resposta a uma determinada exposição ou intervenção por algum estímulo [84]. Ou seja, um biomarcador é um indicador de um fenômeno biológico, que pode ser usado para identificar esse fenômeno ou até mesmo prever o risco de ocorrência dele.

Entende-se que a infecção pelo SARS-CoV-2, assim como outras doenças infecciosas ocasionam respostas imunes desreguladas e em excesso e, no caso da COVID-19, acabam por contribuir para o agravamento da síndrome respiratória aguda promovida pela infecção. A descrição detalhada do perfil imunológico de pacientes em estado crítico sugere hiperativação de vias imunológicas humorais, como por exemplo a IL-6 e a apontando como um mediador crítico para insuficiência respiratória e falência múltipla dos órgãos[85]. A IL-6 é uma potente citocina pró-inflamatória que é liberada após resposta imunológica devida a lesões de tecidos celulares, estimulando assim inúmeras respostas inatas e adaptativas[86]. A expressão elevada de IL-6 pode levar a inúmeras doenças inflamatórias crônicas, o que é observado na COVID-19[87]. Esses episódios inflamatórios podem ser tratados por meio de terapia com anticorpos monoclonais direcionados para o receptor de IL-6 (IL-6R) e sua forma solúvel (Sil-6R) [88]. Dessa forma, entende-se como importante a compreensão do processo de homeostase que regula a resposta imunológica através do IL-6[88,89].

A síntese de IL-6 ativa resposta imune aguda após lesões em tecidos muitas vezes causadas por infecções, e isso induz a diferenciação de células B ativas, células T CD4+ ainda não maduras e plasmócitos produtores de anticorpos[90]. Ela também estimula células hepáticas a produzir proteínas de fase aguda como fibrinogênio, proteína C reativa e hepcidina, sendo que com a elevação dos níveis dessas proteínas, um sinal de estresse emergencial é induzido, estimulando a hematopoiese[90]. Os megacariócitos maduros na medula óssea são responsáveis pela produção de plaquetas, e sabe-se que uma variedade de citocinas pode estimular sua produção. Dessa forma, informações de contagem de plaquetas e linfócitos podem funcionar como indicadores de inflamações e infecções[91].

A proteína C reativa funciona como um marcador de inflamação, sendo regulada por IL-6[90]. Considera-se hoje a IL-6 como um dos principais biomarcadores na COVID-19. Uma metanálise de 2020 realizada com as concentrações médias de IL-6 apontou níveis até 2,9 maiores em pacientes em estado grave do que pacientes sem maiores complicações [92].

Outro biomarcador importante para a intervenção na COVID-19 é o dímero-D, sendo também chamado de D-dímero[93]. Ele é produzido por meio da degradação da fibrina. O sistema fibrinolítico quebra a malha de fibrina após a formação do coágulo, e o dímero-D (dois fragmentos D de fibrina) é formado através da ativação da enzima plamina[93]. Isso acaba indicando a presença de uma fibrina desmontada na corrente sanguínea e o dímero-D representando a ativação dos sistemas de coagulação de fibrinólise[93]. Sua dosagem é comumente usada em diagnósticos ou contra diagnósticos de casos trombolíticos[94]. Observou-se na COVID-19 um aumento considerável de eventos tromboembólicos, o que acabava por causar maiores adversidades no quadro clínico de pacientes já acometidos por uma manifestação mais grave da doença[95]. Estudos tem relatado um aumento das concentrações de dímero-D e fibrinogênio nos estágios iniciais da COVID-19[93]. Os resultados combinados de uma metanálise que avaliou as concentrações de dímero-D mostraram que suas concentrações foram significativamente maiores em pacientes com COVID-19 grave se comparados com pacientes leves[96].

Além disso, observou-se que doenças subjacentes tais como diabetes, derrames ou até mesmo situações como gravidez podem contribuir para o aumento dos níveis de dímero-D pacientes com COVID-19[97]. Vários estudos têm mostrado que a COVID-19 pode aumentar a predisposição dos pacientes a trombose, tanto em artérias quanto veias[94,98,99]. Portanto, os riscos de pacientes COVID-19 positivos são compartilhados com outras doenças de caráter trombolítico, como a trombose venosa profunda e o trombolismo venoso[99]. A

coagulopatia associada à COVID-19 pode ser tratada seguindo a prática estabelecida para o uso de anticoagulantes em casos de trombose, contudo, apesar de sinais de coagulação, tais como dor e inchaço, falta de ar ou dor no tórax, e níveis elevados de dímero-D estejam funcionando como indicadores de mortalidade, o uso de anticoagulantes mais fortes não é recomendado, pois o uso incorreto poderia levar a hemorragias graves. Medicamentos com heparina e enoxaparina foram adotados na clínica e ajustados conforme os níveis de dímero D[100,101].

Muitos polimorfismos genéticos estão sendo explorados nos desfechos clínicos na COVID-19, tanto no prognóstico quanto na suscetibilidade. Muitos trabalhos apontaram como promissores estudos de associação genética envolvendo polimorfismos em genes do sistema renina-angiotensina, mais precisamente nos genes *ACE1* e *ACE2*, e polimorfismos no gene *TMPRSS2*, já que estariam intimamente ligados ao processo infecioso[74,102,103].

Estudos sugerem que o *ACE2* possui papel importante no processo inflamatório, já que pequenos aumentos dos níveis de angiotensina II podem causar a liberação de mediadores inflamatórios, tais como TGF-beta e a IL-6[82,104]. Uma isoforma do *ACE2* foi encontrada em estudos de expressão de *ACE2* induzida por interferon pela infecção por SARS-CoV-2, chamada de *ACE2* truncado, ou *dACE2*[105]. Dessa forma, a variabilidade induzida por interferon nos níveis de expressão de *ACE2* pode ser importante para o desfecho clínico da COVID-19.[105]

Genes envolvidos em vias de interferon, como *IFITM3*, por exemplo, também foram propostos como candidatos, além de diversos loci já estudados em doenças infeciosas, como o fator de necrose tumoral (TNF- ***α***), diversas interleucinas, polimorfismos em genes codificadores de proteases diversas que possam atuar na clivagem do complexo vírus-receptor além da TMPRSS2, como o *CTSL*, *FURIN, IFNL3, IFNL4,* e polimorfismos no complexo *HLA*[74,106–108]*.* O gene *FURIN* foi implicado na infecção da COVID-19 e tem sido associado com maiores riscos de contágio[109,110]. Vários polimorfismos já foram descritos no gene *FURIN.* Um deles é o rs17514846, um polimorfismo de nucleotídeo único (SNP) que está relacionado a níveis maiores de expressão em células do endotélio vascular e já foi associado a síndromes metabólicas[111]. Um SNP no promotor do *FURIN*, o rs4932178, leva a um aumento da atividade transcricional considerável do gene e tem sido associado ao aumento do risco do desenvolvimento da hepatite B crônica[112]. Dessa forma, ele se torna um candidato para estudos de associação envolvendo a infecção por SARS-CoV-2 e o

prognóstico da COVID-19, dada o aumento da transcrição da furina e sua participação efetiva na interação vírus-hospedeiro.

Alguns estudos de associação ao longo do genoma (GWAS) tem levantado candidatos interessantes para estudos de associação genética e estudos funcionais na COVID-19, e um candidato que chamou bastante atenção é o lócus ABO[113] (Figura 3).

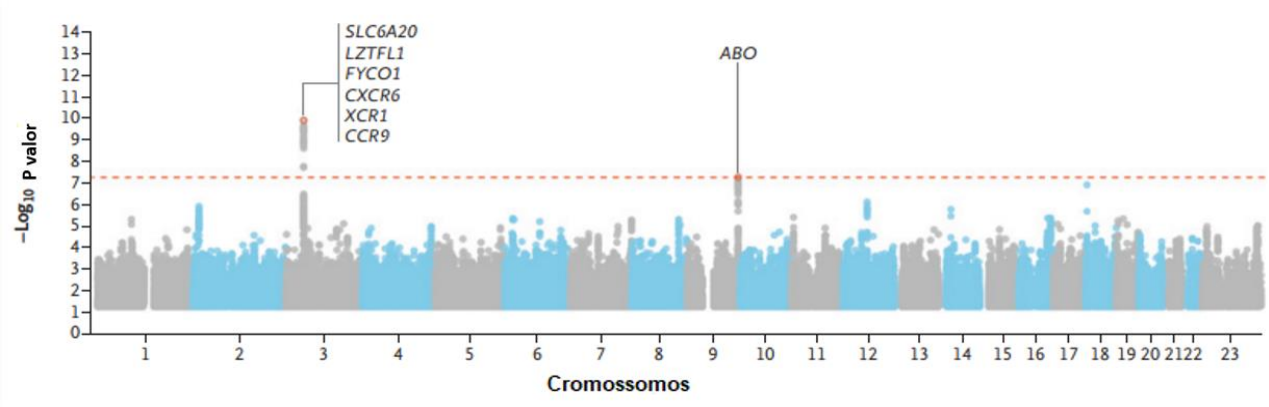


Figura 3 Adaptado de Ellinghaus 2020. GWAS com destaque de dois loci com significância genômica para a severidade da COVID-19 envolvendo insuficiência respiratória.

De acordo com a literatura, indivíduos portadores do tipo sanguíneo O possuem níveis maiores de IL-6 do que os demais tipos, o que sugere vantagens do tipo O na manutenção do papel do *ACE2* no sistema renina-angiotensina[114]. Dessa forma, poderia ocasionar em uma diminuição dos riscos do desenvolvimento de sintomas mais graves na COVID-19. Por outro lado, temos observado o alelo A no lócus *ABO* estando associado ao risco aumentado em doenças cardiovasculares, risco esse relatado em vários estudos[115,116]. Em relação a COVID-19, portadores do alelo A tem demostrado maior susceptibilidade a infecção pelo vírus e ao desenvolvimento de sintomas graves, internações e óbito[117]. Embora o tipo sanguíneo e o histórico de doenças cardiovasculares sejam indicadores de gravidade na COVID-19, eles não podem ser relacionados diretamente ao risco de infecção por SARS-CoV-2, ou até mesmo utilizados de forma determinística ao se tratar do prognóstico. Estudos relacionam o sistema ABO a trombose, onde indivíduos “não O” apresentariam maior risco.[118–120].

Dentre os GWAS já publicados, o lócus 3p21.31 tem chamado bastante atenção, além do sistema ABO[121–123] (Figura 3). Este lócus abrange 6 genes: *SLC6A20*, *CCR9*, *CXCR6*, *XCR1*, *FYCO1* e *LZTFL1*. O *SLC6A20* codifica uma proteína transportadora de solutos, sendo bastante expressa nos rins. Mutações nesse gene podem levar a aminoglicinúria,

que é um distúrbio do transporte tubular renal que afeta a reabsorção de glicina e dos aminoácidos prolina e hidroxiprolina[124,125]. O gene *CCR9* (C-C motif chemokine receptor 9) codifica um receptor de quimiocinas beta. Ele é expresso em linfócitos T do intestino delgado e do cólon, e sua interação com outras quimiocinas contribui para o direcionamento de linfócitos para o intestino delgado, conferindo assim um papel na resposta imune no trato intestinal[126]. O *CXCR6* (C-X-C motif quimiocina receptor 6) codifica uma proteína receptora de quimiocinas e é expressa preferencialmente em células de memória. Ele é bastante estudado na infecção pelo virus HIV, já que funciona como receptor de membrada para a entrada do vírus na celula[127,128]. O gene *XCR1* (X-C motif chemokine receptor 1) codifica um receptor de quimiocina acoplados a proteína G. Essa proteína possui envolvimento com os níveis intracelulares de íons de cálcio. Ele é expresso por células dendríticas no intestino e contribuem para a imunidade contra infecções virais[129]. O gene *FYCO1* (FYVE and coiled-coil domain autophagy adaptor 1) codifica uma proteína adaptadora implicada no transporte de microtúbulos de autofagossomos[130]. Mutações nesse gene estão associados a miosite por corpos de inclusão, uma doença caracterizada por se manifestar tardiamente e por provocar fraqueza muscular progressiva e níveis séricos elevados de creatinoquinase (CK)[131]. O gene *LZTFL1* (Leucine zipper transcription factor like 1) que codifica uma proteína expressa no citoplasma. Este gene já está sendo replicado em estudos de associação como gene candidato no prognóstico da COVID-19 e tem apresentado resultados significativos[132–134].

O desfecho clínico é multifatorial e inúmeros fatores estão relacionados a evolução da doença. A busca por biomarcadores desfechos clínicos e a suscetibilidade à COVID-19 é de extrema importância para que possamos desenvolver métodos de intervenção adequados e que auxilie no controle da pandemia. Alvos terapêuticos podem ser desenvolvidos tendo como base marcadores funcionais, além de intervenções sociais serem organizadas baseadas na dinâmica populacional de loci de risco dependendo da população de interesse e das realidades genéticas locais. A resposta a possíveis futuros medicamentos e abordagens terapêuticas também dependem desse background genético populacional, e a compreensão da genética do hospedeiro é fundamental para o combate de qualquer doença, mesmo infecciosa.

Sendo assim, esta tese tem como objetivo principal investigar a contribuição de possíveis biomarcadores genéticos na evolução do quadro clínico da COVID-19, além de buscar entender a influência de tais marcadores no aumento ou diminuição do risco à infecção pelo SARS-CoV-2.

CAPÍTULO 1

Neste capítulo, conduzimos uma revisão sistemática seguindo as orientações presentes na recomendação PRISMA (https://www.prisma-statement.org/). O Objetivo foi levantarmos os estudos que abordaram genética do hospedeiro e desfechos clínicos da COVID-19 até o final de 2020.

Este trabalho foi aceito em 22 de julho de 2021 e publicado em 02 de agosto de 2021 no periódico Reviews in Medical Virology (doi: 10.1002/rmv.2283).

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

DOI: 10.1002/rmv.2283

WILEY

REVIEW

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

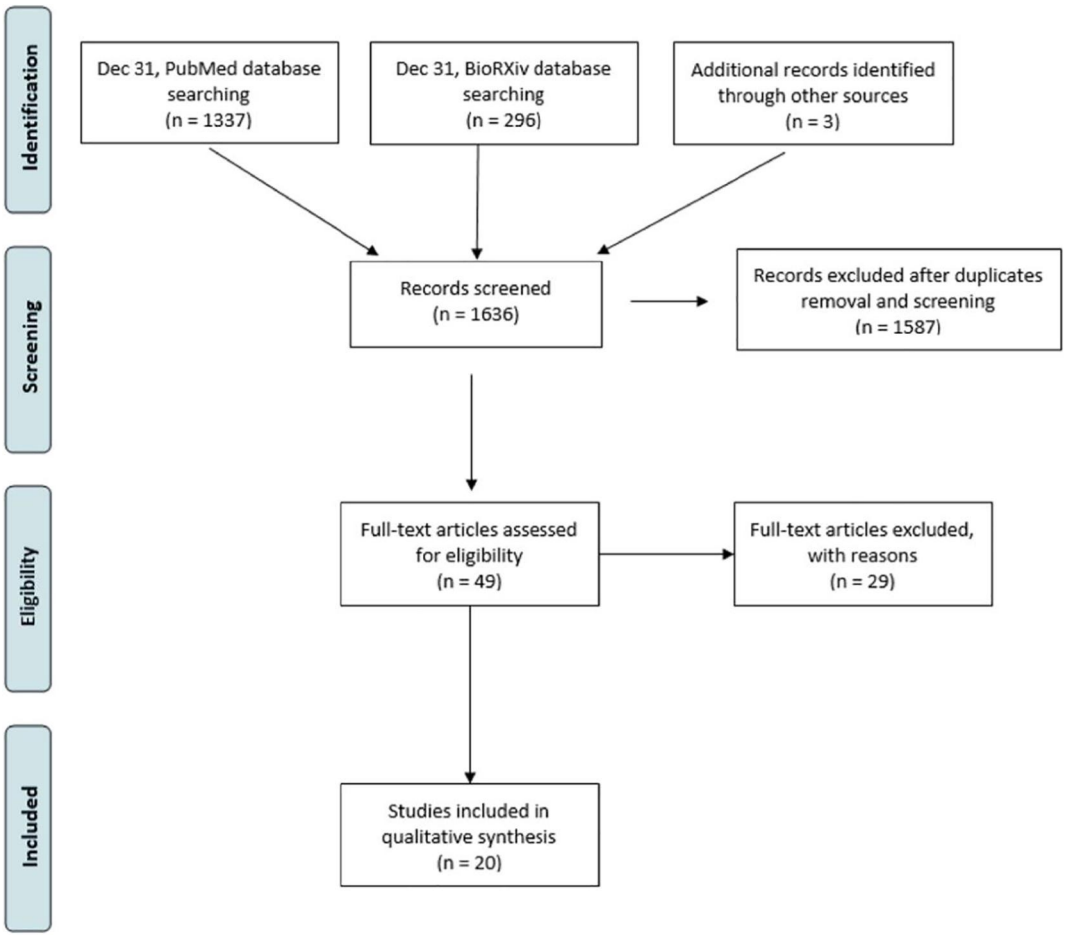


FIGURE 1 PRISMA flow diagram

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

CAPÍTULO 2

Neste capítulo, conduzimos uma revisão sistemática seguida de metanálise com base nas orientações presentes na recomendação PRISMA. Nosso objetivo foi investigar a contribuição de marcadores nos genes *ACE1*, *IFITM3*, *FURIM* e *TNF-* ***α*** no prognóstico da COVID-19. Este trabalho foi publicado no periódico Frontiers in Genetics em 01 de abril de 2022 (DOI: 10.3389/fgene.2022.775246).

frontiers | Frontiers in Genetics

SYSTEMATIC REVIEW
published: 01 April 2022
doi: 10.3389/fgene.2022.775246

IFITM3, *FURIN*, *ACE1*, and *TNF-α* Genetic Association With COVID-19 Outcomes: Systematic Review and Meta-Analysis

*João Locke Ferreira de Araújo, Diego Menezes, Renato Santana de Aguiar and Renan Pedra de Souza**

Grupo de Pesquisa em Bioestatística e Epidemiologia Molecular, Laboratório de Biologia Integrativa, Programa de Pós Graduação em Genética, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

Human polymorphisms may contribute to SARS-CoV-2 infection susceptibility and COVID-19 outcomes (asymptomatic presentation, severe COVID-19, death). We aimed to evaluate the association of *IFITM3*, *FURIN*, *ACE1*, and *TNF-α* genetic variants with both phenotypes using meta-analysis. The bibliographic search was conducted on the PubMed and Scielo databases covering reports published until February 8, 2022. Two independent researchers examined the study quality using the Q-Genie tool. Using the Mantel–Haenszel weighted means method, odds ratios were combined under both fixed- and random-effect models. Twenty-seven studies were included in the systematic review (five with *IFITM3*, two with *Furin*, three with *TNF-α*, and 17 with *ACE1*) and 22 in the meta-analysis (*IFITM3* $n = 3$, *TNF-α*, and *ACE1* $n = 16$). Meta-analysis indicated no association of 1) *ACE1* rs4646994 and susceptibility, 2) *ACE1* rs4646994 and asymptomatic COVID-19, 3) *IFITM3* rs12252 and ICU hospitalization, and 4) *TNF-α* rs1800629 and death. On the other hand, significant results were found for *ACE1* rs4646994 association with COVID-19 severity (11 studies, 692 severe cases, and 1,433 nonsevere controls). The *ACE1* rs4646994 deletion allele showed increased odds for severe manifestation (OR: 1.45; 95% CI: 1.26–1.66). The homozygous deletion was a risk factor (OR: 1.49, 95% CI: 1.22–1.83), while homozygous insertion presented a protective effect (OR: 0.57, 95% CI: 0.45–0.74). Further reports are needed to verify this effect on populations with different ethnic backgrounds.

Systematic Review Registration: https://www.crd.york.ac.uk/prosperodisplay_record.php?ID=CRD42021268578, identifier CRD42021268578

Keywords: polymorphism, genetic association study, candidate genes, transposable elements, biomarkers, host genetics

OPEN ACCESS

Edited by:
José M. Álvarez-Castro, University of Santiago de Compostela, Spain

Reviewed by:
Martha Guevara-Cruz, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (INCMNSZ), Mexico
Gagandeep Kaur, University of Rochester, United States

****Correspondence:***
Renan Pedra de Souza
renanrps@ufmg.br

Specialty section:
This article was submitted to Applied Genetic Epidemiology, a section of the journal Frontiers in Genetics

Received: *13 September 2021*
Accepted: *11 February 2022*
Published: *01 April 2022*

Citation:
de Araújo JLF, Menezes D, Aguiar RSd and Souza RPd (2022) IFITM3, FURIN, ACE1, and TNF-α Genetic Association With COVID-19 Outcomes: Systematic Review and Meta-Analysis. Front. Genet. 13:775246. doi: 10.3389/fgene.2022.775246

INTRODUCTION

Coronavirus disease 2019 (COVID-19) clinical presentation is heterogeneous, ranging from entirely asymptomatic up to severe cases and death. Another level of heterogeneity is observed regarding persistent symptoms: one study has estimated that the median proportion of individuals who experienced at least one persistent symptom was 73% (Nasserie et al., 2021). Uncovering biomarkers

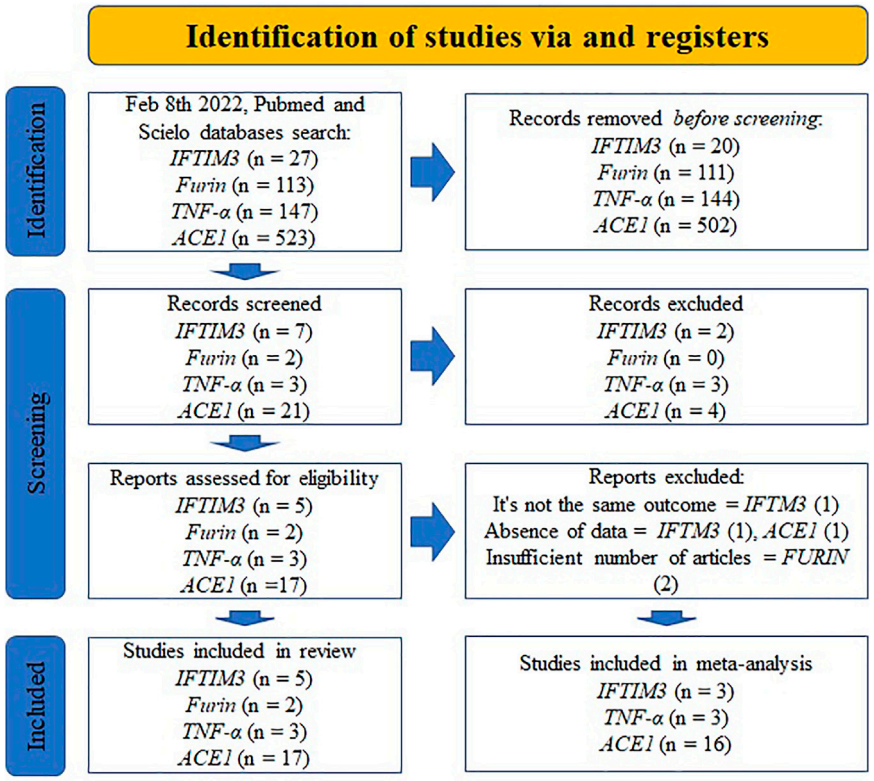


FIGURE 1 | Study selection using Preferred Reporting Items for Systematic reviews and Meta-analysis (PRISMA) guidelines (16).

linking patients with distinct prognosis subgroups would be beneficial. Different strategies have been employed to uncover molecular markers predicting odds for better prognosis and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection susceptibility. Proteins, lipids, and metabolites have already been examined (Praissman and Wells, 2021; Samprathi and Jayashree, 2021). Genetic variability has been shown to be a valuable source for biomarker research. COVID-19 prognosis and infection susceptibility are multifactorial traits determined by the complex interaction of environmental factors and multiple genes. Thus, significant single-gene results may lead to substantial predictors such as the C–C chemokine receptor type five (*CCR5*) gene association with HIV susceptibility and prognosis (Liu et al., 2012), or ABO blood type and dengue severity (Hashan et al., 2021).

Genetic association studies can be designed within prespecified genes of interest (candidate gene approach) or with a broader strategy characterizing diversity across large genomic areas (genome-wide association studies, whole exome and genome sequencing). Angiotensin-converting enzyme 2 (*ACE2*), transmembrane serine protease 2 (*TMPRSS2*), human leukocyte antigen (*HLA*), interferon-induced transmembrane protein 3 (*IFITM3*), tumor necrosis factor-alpha (*TNF-α*), *FURIN*, and angiotensin I-converting enzyme (*ACE1*) were the most studied genes using the candidate gene approach in 2020 (Araújo et al., 2021). They all present strong biological plausibility since they act on viral cell entry or human immune response to SARS-CoV-2.

Findings from single association studies must always be considered carefully because of the likelihood of producing spurious outcomes (Sullivan, 2007). Replication is essential before considering using genetic markers in the clinical setting. Although that has been proved hard, inconsistency frequently can be attributed to shortfalls in study design, implementation, and interpretation, with inadequately powered sample groups being of significant concern (Hattersley and McCarthy, 2005). A systematic meta-analytic approach may support estimating population-wide effects of genetic risk factors in human diseases (Ioannidis et al., 2001). The PROSPERO (Moher et al., 2014) database, indicating protocols for systematic reviews, has already been presented for *HLA* (CRD42021251670) (Deb et al., 2022), *ACE2*, and *TMPRSS2* (CRD42021229963) contribution with COVID-19 outcomes. Therefore, we focused our systematic review on *IFITM3*, *FURIN*, *ACE1*, and *TNF-α* genetic variants and their association with COVID-19 susceptibility and prognosis to reduce unnecessary duplication.

IFITM3 (MIM 605579; 11p15.5) is a protein-coding gene that disturbs cell entry by inhibiting viral fusion with cholesterol-depleted endosomes (Amini-Bavil-Olyaee et al., 2013); a mechanism also described during SARS-CoV-2 infection (Prelli Bozzo et al., 2021). The *IFITM3* rs12252 polymorphism has been associated with influenza severity (Prabhu et al., 2018). The *TNF* (MIM 191160; 6p21.33) gene encodes a multifunctional proinflammatory cytokine. Although TNF-α is not as relevant as

TABLE 1 | Association studies of *ACE1* rs4646994 (Alu 287 pb) with coronavirus disease 2019 (COVID-19) susceptibility included in the systematic review.

Year	Author	Sample					Control		Case	
		Date	Place	Ethnic background	Size	Male n(%)	n	Criteria	n	Criteria
2020	Gòmez	—	Spain	Caucasian (Asturias)	740	373 (0.50)	536	Healthy population	204	COVID-19 positive
2021	Akbari	2020	Iran	—	182	105 (0.57)	91	Unaffected individuals without a history of exposure to COVID-19 cases	91	COVID-19 positive
	Aladag	May/2020	Turkey	—	412	—	300	General population	112	COVID-19 positive
	Annunziata	March–April/20	Italy	Southern Italians	39	—	19	Healthy subjects	20	COVID-19 positive
	Hubacek	March–June/2020	Czech Republic	—	2,989	–(0.54)	2,579	General population	408	COVID-19 positive
	Kouhpayeh	May–September/2020	Iran	—	520	276 (0.55)	258	Healthy subjects with negative PCR and clinical diagnostic criteria	244	COVID-19 positive
	Mahmood	October–December/2020	Iraq	—	195	–(0.50)	96	Healthy subjects with negative serological test	99	COVID-19 positive
	Mir	September/2020–April/2021	Saudi Arabia	—	267	185 (0.69)	150	Healthy subjects	117	COVID-19 positive
	Möhlendick	March–September/2020	Germany	—	550	323 (0.59)	253	Patients with COVID-19 symptoms with negative PCR	297	COVID-19 positive
	Saad	—	Lebanon	Lebanese	387	195 (0.50)	155	Participants with negative PCR	232	COVID-19 positive
2022	Gong	January–March/2020	China	—	862	—	441	Healthy subjects	421	COVID-19 positive
	Papadopoulou	March–June/2020	Greece	Caucasian (Greek)	389	—	316	Blood product donors and volunteer healthcare workers	73	COVID-19 positive

interleukin-6 on the cytokine storm presented in severe COVID-19 patients (Karki and Kanneganti, 2021), anti-TNF-α drug repositioning for COVID-19 has been proposed (Stebbing et al., 2020). FURIN is coded by the *FURIN* (MIM 136950; 15q26.1) gene. It regulates constitutive exocytic and endocytic pathways and has a central role in SARS-CoV-2 transmission (Peacock et al., 2021). The *ACE1* (MIM 106180; 17q23.3) gene produces a protein related to blood pressure regulation and electrolyte balance, and ACE1/ACE2 balance has been suggested to play a pivotal role in the pathobiology and treatment of COVID-19 (Sriram and Insel, 2020). The *ACE1* rs4646994 variant is a 287-bp Alu repeat insertion/deletion (indel) on intron 16 known to alter ACE-1 levels and influence several clinical traits (Castellon and Hamdi, 2007). Here, we present the result of a systematic review and, whenever possible, a meta-analysis of *IFITM3*, *FURIN*, *ACE1*, and *TNF-α* genetic association with susceptibility to SARS-CoV-2 infection and COVID-19 severity.

MATERIALS AND METHODS

The systematic review protocol was submitted to PROSPERO (CRD42021268578). Preferred Reporting Items for Systematic reviews and Meta-analysis (PRISMA) was adopted as a guideline for reporting this systematic review (Page et al., 2021). The study selection was carried out in three phases: identification, screening, and eligibility. Search on the PubMed and Scielo databases led to article identification. The PECO question for prognosis was Participants (P) = subjects with COVID-19, Exposition (E) = minor alleles, Control (C) = major alleles of genetic variants, and Outcomes (O) = COVID-19 severity (asymptomatic or severe presentation); while the PECO question for susceptibility was P = overall population, E = minor alleles, C = major alleles of genetic variants, and Outcomes (O) = COVID-19 positive diagnosis. The bibliographic search included all studies published until February 8, 2022, with no language restriction, using the search arguments listed in **Supplementary Material SI**. Two independent researchers conducted article screening. Inclusion criteria were primary articles covering genetic association of COVID-19 susceptibility or prognosis with *IFITM3*, *FURIN*, *ACE1*, and *TNF-α* variants, comprising four separate searches. Exclusion criteria were review articles or primary articles evaluating the association of COVID-19 susceptibility or prognosis with other genes.

We assessed study quality using the Q-Genie tool (Sohani et al., 2016) performed by two independent researchers. This instrument contains 11 questions to be marked on a seven-point Likert scale examining 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 nontechnical classification of tested genetic variants (e.g., genotyping call rates, blinded experiments), classification of the outcome (e.g., sampling

TABLE 2 | Association studies of *ACE1* rs4646994 (Alu 287 pb) with COVID-19 prognosis included in the systematic review.

Phenotype	Year	Author	Sample					Control		Case	
			Date	Place	Ethnic background	Size	Male n(%)	n	Criteria	n	Criteria
Asymptomatic × symptomatic	2021	Cafiero	—	Italy	—	104	58 (0.56)	50	Asymptomatic	54	Symptomatic (x-ray imaging)
		Hubacek	March–June/2020	Czech Republic	—	408	–(0.55)	163	Asymptomatic	245	Symptomatic (no hospitalization)
		Gunal	April–July/2020	Turkey	—	60	—	30	Asymptomatic	30	Severe (RR ≥30/min; SpO_2 ≤93%; PaO_2/FiO_2 ≤300 mmHg; mechanical ventilation or ICU)
Nonsevere × severe	2020	Gòmez	—	Spain	Caucasian (Asturias)	204	125 (0.61)	137	Mild (hospitalized, nonsevere)	67	Severe (hospitalized, mechanical ventilation and/or ICU)
	2021	Akbari	2020	Iran	—	91	53 (0.58)	54	Hospitalized, non-ICU	37	Hospitalized, ICU
		Aladag	May/2020	Turkey	—	65	-	53	Nonsevere	12	Severe (fever or suspected respiratory infection, plus one of the following: RR >30/min; severe respiratory distress; or SpO_2 ≤93%)
		Çelik	—	Turkey	—	154	78 (0.50)	119	Mild (outpatients) and moderate (hospitalized nonsevere)	35	Severe (RR ≥30/min; SpO_2 ≤93%; PaO_2/FiO_2 ≤300 mmHg; mechanical ventilation or ICU)
		Gunal	April–July/2020	Turkey	—	90	-	60	Asymptomatic and mild	30	Severe (RR ≥30/min; SpO_2 ≤93%; PaO_2/FiO_2 ≤300 mmHg; mechanical ventilation or ICU)
		Kouhpayeh	May–September/2020	Iran	—	258	144 (0.56)	106	Nonsevere	152	Severe (fever or suspected respiratory infection, plus one of the following: RR >30/min; severe respiratory distress; or SpO_2 ≤93%)
		Mahmood	October–December/2020	Iraq	—	99	–(0.51)	68	Mild (with symptoms of pneumonia and no signs of severe pneumonia)	31	Severe (severe respiratory distress, RR ≥30 breaths/min or SpO_2 ≤ 93%)
		Möhlendick	March–September/2020	Germany	—	251	176 (0.59)	207	Mild and hospitalized (non-ICU)	44	Severe (hospitalized, mechanical ventilation and/or ICU)
		Saad	-	Lebanon	Lebanese	223	123 (0.55)	162	Mild and moderate	61	Severe (lung infiltrates on chest x-ray or CT scan and SpO_2 <94% who required hospitalization with essential oxygen therapy or mechanical ventilation)
		Verma	August–September/2020	India	India	269	174 (0.65)	149	Mild (RR <24/min, SpO_2 >94%)	120	Severe (pneumonia with RR > 30/min; severe respiratory distress; or SpO_2 ≤93%)
	2022	Gong	January–March/2020	China	—	421	—	318	Mild and moderate	103	Severe
		Papadopoulou	March–June/2020	Greece	Caucasian (Greek)	81	43 (0.53)	29	Mild and moderate (with symptoms of pneumonia and no signs of severe pneumonia)	52	Severe or critical (fever or suspected respiratory infection, plus one of the following: RR >30/min; severe respiratory distress; or SpO_2 ≤93%)
Alive × dead	2021	Mir	September/2020–April/2021	Saudi Arabia	—	117	85 (0.73)	74	Alive	43	Dead
		Möhlendick	March–September/2020	Germany	—	297	176 (0.59)	251	Mild, hospitalized (non-ICU) and severe	46	Dead

Note. RR, respiratory rate; ICU, intensive care unit; SpO_2, oxygen saturation; PaO_2/FiO_2, arterial oxygen pressure/fraction of inspired oxygen; CT, computerized tomography.

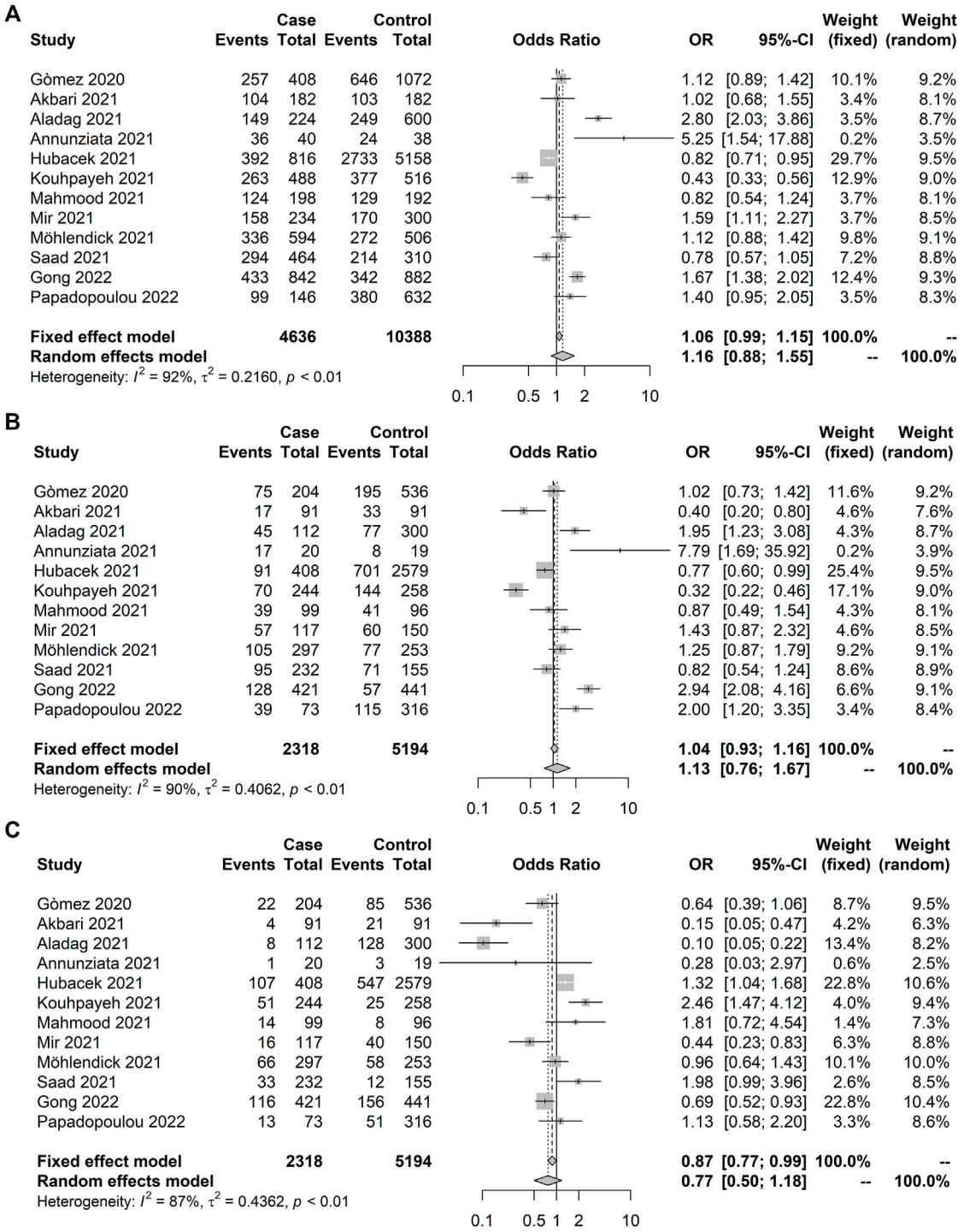


FIGURE 2 | Forest plot illustrating *ACE1* rs4646994 (Alu 287 pb) association with coronavirus disease 2019 (COVID-19) susceptibility. No significant results were observed. Case and control definitions are presented in **Table 1**. **(A)** C allele association. **(B)** C recessive model. **(C)** T recessive model.

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. Proposed cutoffs for understanding are ≤35 poor, > 35 moderate, and ≥45 good quality, with the total score ranging from 7 to 77 points.

Meta-analysis was conducted whenever three or more studies were included for the same polymorphism and outcome. We

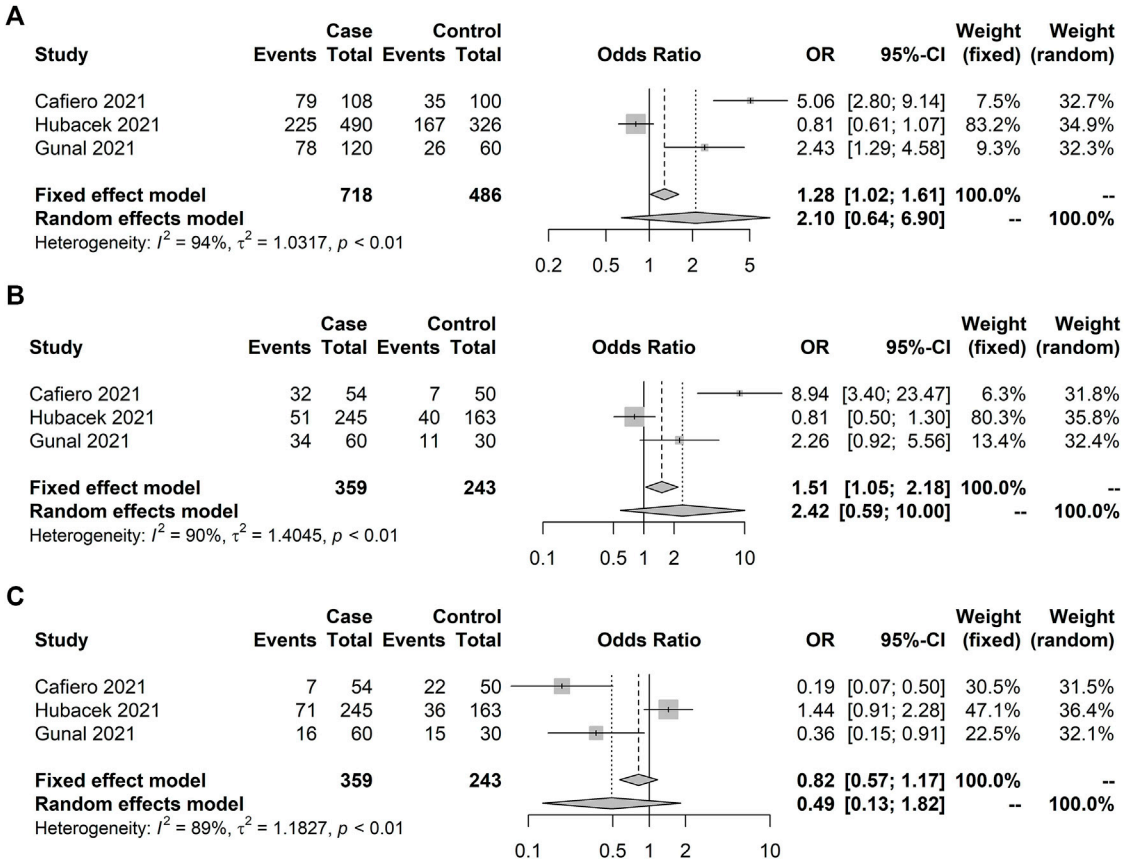


FIGURE 3 | Forest plot illustrating *ACE1* rs4646994 (Alu 287 pb) association with symptom presence (asymptomatic × symptomatic). No significant allelic and genotypic effects were observed under the random model. Case and control definitions are presented in **Table 2**. **(A)** D-allele model. **(B)** D recessive model. **(C)** I recessive model.

carried out single meta-analyses for each polymorphism considering allelic and genotypic effects (under both allele recessive model assumptions). Heterogeneity between studies was assessed using the chi-square test. We used the *metabin* function coded on *meta* package in R (version 4.1.0) (R Core Team, 2014) to estimate overall odds ratios (ORs) and its 95% confidence interval (CI). Original ORs were combined using the Mantel–Haenszel weighted means method under both fixed- and random-effect models. The significance level was set at 0.05.

RESULTS

Twenty-seven studies were included in the systematic review: five with *IFITM3* (Zhang et al., 2020; Alghamdi et al., 2021; Cuesta-Llavona et al., 2021; Gómez et al., 2021; Schönfelder et al., 2021), two with *Furin* (Latini et al., 2020; Torre-Fuentes et al., 2021), three with *TNF-α* (Saleh et al., 2020; Fishchuk et al., 2021; Heidari Nia et al., 2021), and 17 with *ACE1* (Gómez et al., 2020; Aladag et al., 2021; Annunziata et al., 2021; Cafiero et al., 2021; Gunal et al., 2021; Hubacek et al., 2021; Karakaş Çelik et al., 2021; Kouhpayeh et al., 2021; Mir et al., 2021; Möhlendick et al., 2021; Saad et al., 2021; Verma et al., 2021; Akbari et al., 2022; Gong et al., 2022; Mahmood et al., 2022; Papadopoulou et al., 2022). (**Figure 1**). Inconsistencies in reported frequencies were found in two studies (Gómez et al., 2021; Karakaş Çelik et al., 2021).

All manuscripts but one reached moderate or good quality scores in the Q-Genie analysis (**Supplementary Material S1**). Among the 11 questions, it is clear that all studies had the worst performance for questions number 5 and 10. While question 5 examines reported information regarding how genotyping was conducted (e.g., blinded experiments, batch effects), question 10 evaluated whether genetic relationships among subjects were tested, and sex and ethnicity were stated.

Five meta-analyses were carried out, including 22 studies evaluating three genes (*IFITM3* n = 3, *TNF-α* n = 3, and *ACE1* n = 16). Twelve studies, including 2,318 control subjects and 5,194 COVID-19 positives, evaluated the *ACE1* rs4646994 association with COVID-19 susceptibility (**Table 1**). Significant heterogeneity was observed for all genetic models with no significant association under the random model (**Figure 2**). Similar findings were detected in the meta-analysis of the *ACE1* rs4646994 variant with asymptomatic presentation (**Table 2**), indicating no significant effect pooled from three studies (**Figure 3**). We observed high heterogeneity in the sampling places and reported ethnic backgrounds.

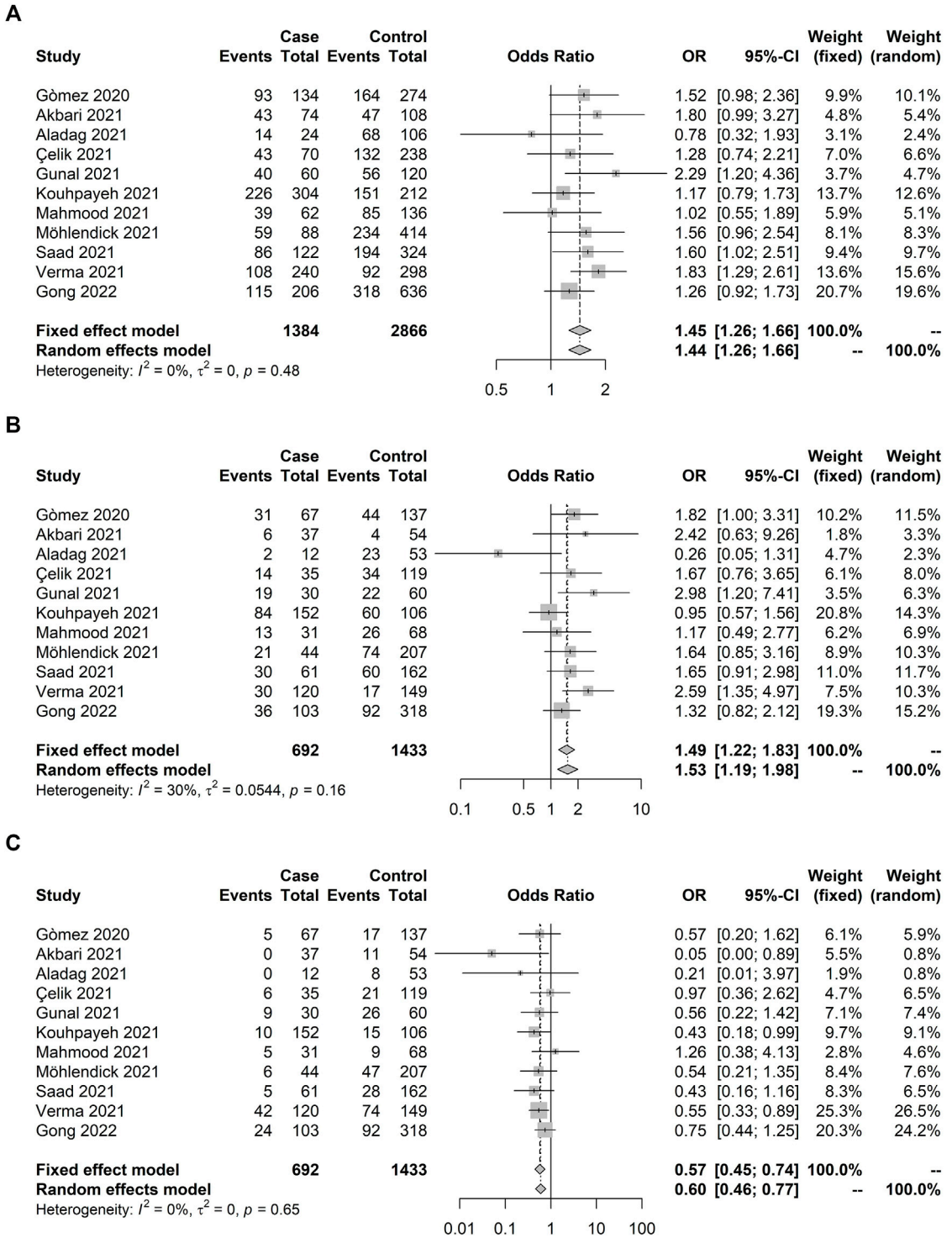


FIGURE 4 | Forest plot illustrating *ACE1* rs4646994 (Alu 287 pb) association with COVID-19 severity (severe × others). Significant allelic and genotypic effects were observed. Case and control definitions are presented in **Table 2**. **(A)** D-allele model. D-allele was associated with increased risk of COVID-19 severity. **(B)** D recessive model. D/D genotype carriers showed increased odds to manifest severe COVID-19 compared with D/I and I/I carriers combined **(C)** I recessive model. I/I genotype carriers showed decreased odds to present severe COVID-19 compared with D/I and D/D carriers combined.

We were able to conduct a meta-analysis investigating whether *ACE1* rs4646994 polymorphism could predict COVID-19 severity. Eleven studies were included reaching a total of 692 individuals with severe COVID-19 and 1,433 with nonsevere manifestation (**Table 2**). The allelic association was observed with increased odds for deletion (D) allele compared with I-allele

TABLE 3 | Association studies of *IFITM3* rs12252 with COVID-19 prognosis included in the systematic review.

Phenotype	Year	Author	Sample					Control		Case	
			Date	Place	Ethnic background	Size	Male n(%)	n	Criteria	n	Criteria
Non-ICU × ICU	2021	Alghamdi	—	Saudi Arabia	Saudi	376	112 (0.56)	210	Hospitalized, non-ICU	166	Hospitalized, ICU
	2021	Cuesta-Llavona	March–December/2020	Spain	Caucasian (Asturias)	484	276 (0.57)	332	Hospitalized, non-ICU	152	Hospitalized, ICU
	2021	Gómez	March–August/2020	Not informed	Caucasian (Asturias)	311	174 (0.56)	230	Hospitalized, non-ICU	81	Hospitalized, ICU
	2021	Schonfelder	March–September/2020	Germany	Caucasian	239	141 (0.59)	164	Outpatients and hospitalized (non-ICU)	75	Hospitalized (ICU or mechanical ventilation) or dead
Alive × dead	2021	Alghamdi	—	Saudi Arabia	Saudi	861	—	784	Alive	77	Dead
	2021	Cuesta-Llavona	March–December/2020	Spain	Caucasian (Asturias)	484	276 (0.57)	114	Alive	38	Dead
Other	2020	Zhang	January–February/2020	China	—	80	33 (0.41)	56	Mild (hospitalized with fever, respiratory symptoms, and pneumonia seen with imaging)	24	Severe (RR ≥30/min; SpO_2 ≤93%; PaO_2/FiO_2 ≤300 mmHg; mechanical ventilation or ICU)
	2021	Alghamdi	—	Saudi Arabia	Saudi	861	—	457	Nonhospitalized	374	Hospitalized

Note. RR, respiratory rate; ICU, intensive care unit; SpO_2, oxygen saturation; PaO_2/FiO_2, arterial oxygen pressure/fraction of inspired oxygen.

(pooled OR: 1.45; 95% CI: 1.26–1.66) (**Figure 4A**). Homozygous deletion (D/D) carriers showed 49% increased odds to present severe COVID-19 compared with heterozygous (D/I) and homozygous insertion allele (I/I) carriers combined (pooled OR: 1.49, 95% CI: 1.22–1.83) (**Figure 4B**). On the other hand, the I/I genotype was protective against severe COVID-19 (pooled OR: 0.57, 95% CI: 0.45–0.74) (**Figure 4C**).

The *IFITM3* rs12252 meta-analysis with severity included three studies totaling 308 individuals admitted to an intensive care unit and 726 who were not admitted (**Table 3**). No significant association was observed under any genetic model (**Figure 5**). Meta-analysis for other outcomes with the *IFITM3* rs12252 could not be conducted. The *TNF-α* rs1800629 association with death was analyzed in three studies (**Table 4**), including 111 subjects who died and 1,095 survivors. No significant association was observed under the random-effect models (**Figure 6**). *FURIN* (**Table 5**) genetic variants had less than three studies; therefore, no meta-analyses were carried out.

DISCUSSION

We conducted a systematic review followed by meta-analysis including studies covering genetic association of COVID-19 susceptibility or prognosis with *IFITM3*, *FURIN*, *ACE1*, and *TNF-α* variants. Four studies included in the meta-analyses did not report the sample collection date, which is of particular interest in COVID-19 studies due to the emergence of variants of concern (VOCs) in the last part of 2020 (Konings et al., 2021). Some VOCs have been associated with higher viral load, worse prognosis, and lethality (Davies et al., 2021; Faria et al., 2021), thus, confounding factors when evaluating genetic effects. Age can also be a confounding factor for COVID-19 association analysis (Fernández Villalobos et al., 2021). Most studies failed to conduct age-corrected estimation or even describe age separately for case and control groups. The same trend was observed for comorbidities (data now shown).

Ancestrality could also contribute to COVID-19 outcomes. Several studies do not present the ethnic background or, at least, the place of birth of the included subjects. Although heterogeneity was seen in parameters associated with ancestrality, the literature fails on genetic background diversity, an issue already raised for genomic data before (Popejoy and Fullerton, 2016). Another literature issue that needs attention is the selective reporting biases leading to the more likely publication of positive findings (Munafò et al., 2009; Sagoo et al., 2009).

We did not find an association of *IFITM3* rs12252 with Covid-19 severity. Our results corroborate the most extensive association study published to date since no significance was reported on any of chromosome 11 loci (Niemi et al., 2021). However, the second evaluated polymorphism, the *ACE1* rs4646994, showed significant effects with homozygous D carriers presenting higher odds of developing severe COVID-19. Several hits on the large arm of chromosome 17 have been previously reported (Niemi et al., 2021), although their genomic location is too far to hypothesize linkage disequilibrium. It is important to note that genome-wide data may find hits on loci that not necessarily are the ones harboring the causative variants because of its experimental design (Spencer et al., 2009). Furthermore, candidate-gene, whole-exome or whole-genome sequencing studies are more suitable in exploring large indel variants.

The *ACE1* rs4646994 has been associated with several clinical phenotypes, including COVID-19 (Castellon and Hamdi, 2007; Li

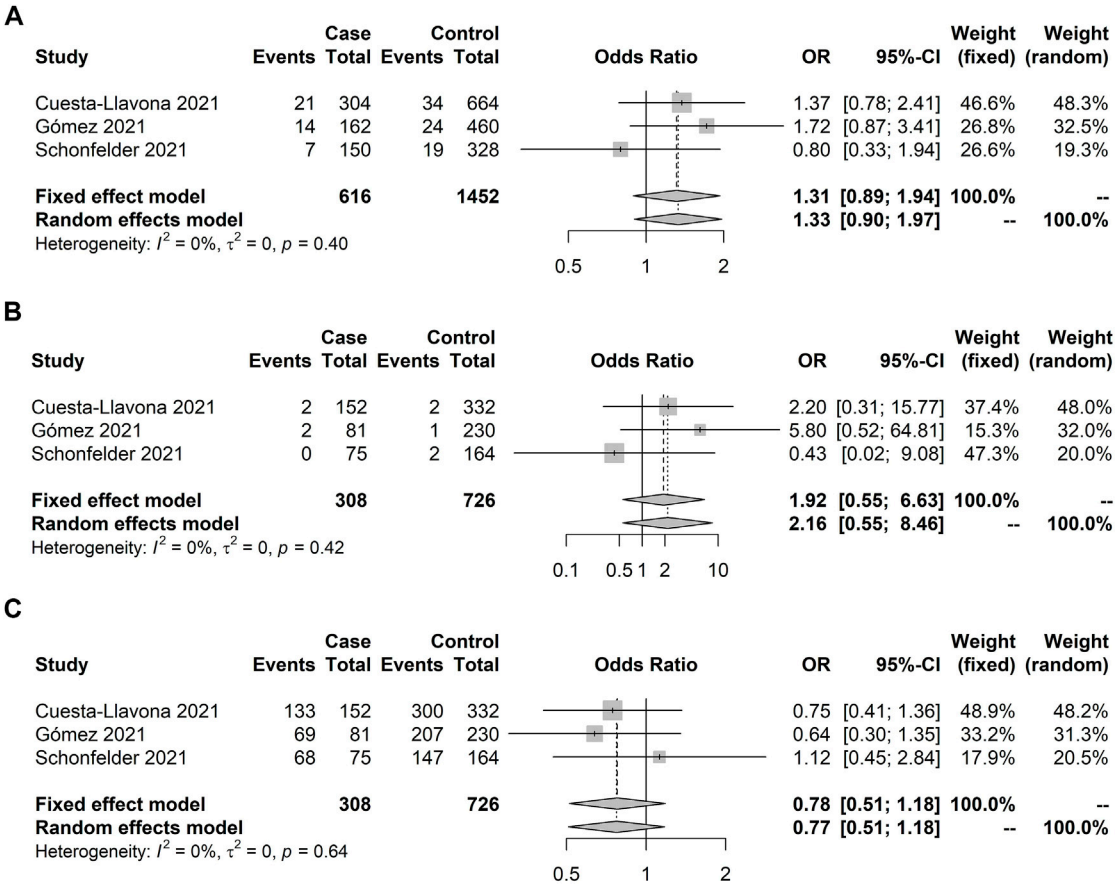


FIGURE 5 | Forest plot illustrating *IFITM3* rs12252 association with severity (non-ICU × ICU). No significant results were observed. Case and control definitions are presented in **Table 3**. **(A)** C allele association. **(B)** C recessive model. **(C)** T recessive model.

TABLE 4 | Association studies of *TNF-α* rs1800629 gene with COVID-19 prognosis or susceptibility included in the systematic review.

Year	*Author*	*Sample*					*Control*		*Case*	
		Date	*Place*	*Ethnic background*	*Size*	*Male n(%)*	*n*	*Criteria*	*n*	*Criteria*
2020	Saleh	April—July/2020	Egypt	—	1,084	600 (0.56)	184	Health care workers	900	COVID-19 positive
					900	-	444	Mild	456	Severe
					900	504 (0.56)	840	Alive	60	Dead
2021	Nia	June/2020—January/2021	Iran	—	550	234 (0.43)	275	COVID-19 negative	275	Hospitalized
					275	112 (0.41)	96	Nonsevere	179	Severe
					275	-	249	Alive	26	Dead
2021	Fishchuk	April–June/2020	Ukraine	—	31	16 (0.50)	25	Alive	6	Dead

et al., 2021). Most previous findings report associations with COVID-19 outcomes on a population level, indicating high variability on allelic frequencies across different populations (Delanghe et al., 2020; Pati et al., 2020; Yamamoto et al., 2020). On a molecular level, expression results indicate increased levels of ACE1 in D-allele carriers (Suehiro et al., 2004) with increased angiotensin II production (Hamdi and Castellon, 2004) and decreased ACE2 protein levels in lung tissue, thereby potentially affecting infectivity by SARS-CoV-2 (Jacobs et al., 2021). Our group has previously indicated that lower ACE2 levels may increase the risk of COVID-19 respiratory distress (Rossi et al., 2021). Although there is a robust biological hypothesis linking *ACE1* rs4646994 with COVID-19, further reports are needed to understand better whether *ACE1* variants could contribute to COVID-19 severity. Moreover, studies are still required to adequately evaluate *IFITM3*, *FURIN*, and *TNF-α* genetic variants' role in COVID-19 susceptibility and outcomes.

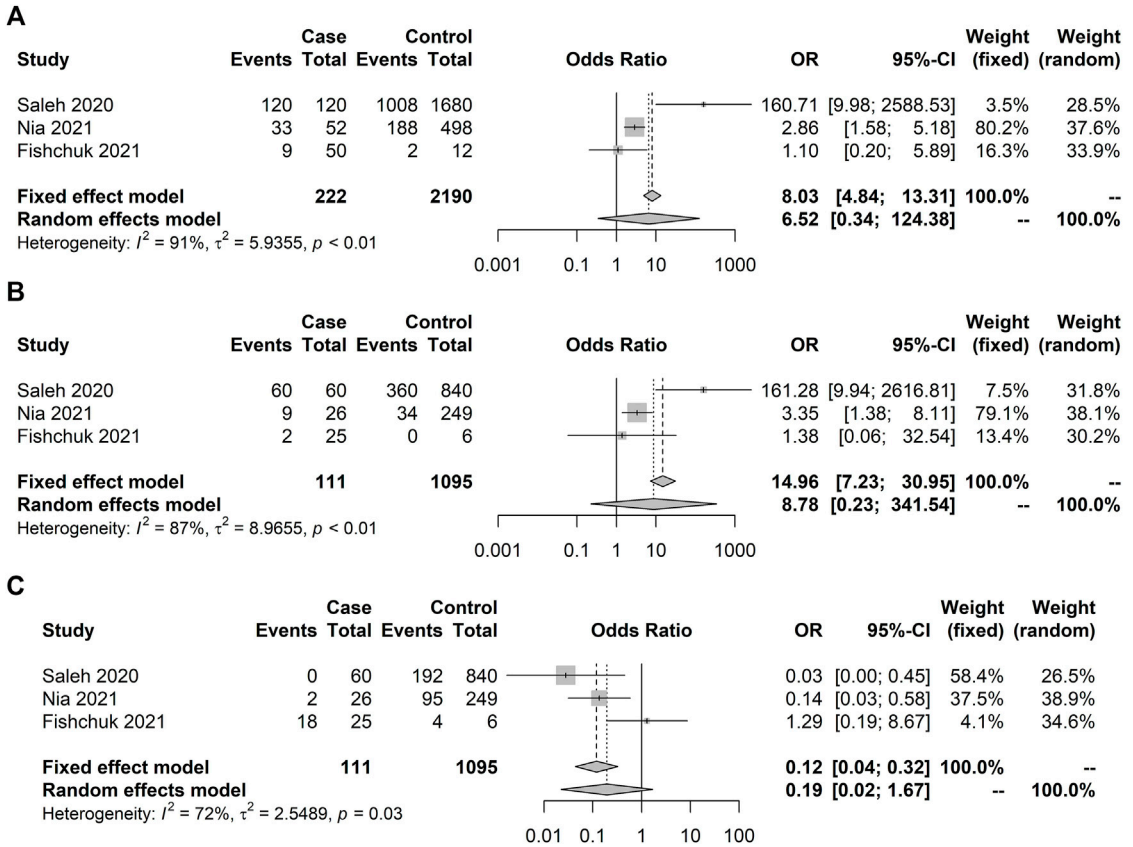


FIGURE 6 | Forest plot illustrating *TNF-α* rs1800629 association with death (alive × dead). No significant results were observed under the random model. Case and control definitions are presented in **Table 5**. **(A)** C allele association. **(B)** C recessive model. **(C)** T recessive model.

TABLE 5 | Association studies of *FURIN* gene with COVID-19 prognosis or susceptibility included in the systematic review.

Year	*Author*	*Sample*					*Control*		*Case*	
		Date	*Place*	*Ethnic background*	*Size*	*Male n(%)*	*n*	*Criteria*	*n*	*Criteria*
2020	Latini	Mar—May/ 2020	Italy	—	—	—	—	Severe (respiratory impairment, requiring noninvasive ventilation)	—	Extremely severe (requiring invasive ventilation and ICU)
					131	82 (0.63)	—	Asymptomatic	—	Severe and extremely severe
2021	Torre-Fuentes	—	Spain	—	120	-	113	COVID-19 negative	7	COVID-19 positive

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

RS wrote the systematic review protocol. JA, DM, and RS conducted the systematic review. JA, RA, and RS drafted the manuscript. All authors revised and approved the final manuscript version.

ACKNOWLEDGMENTS

JA receives a FAPEMIG graduate fellowship. RA and RS are CNPq-Brazil Research Fellows.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.775246/full#supplementary-material

REFERENCES

Akbari, M., Taheri, M., Mehrpoor, G., Eslami, S., Hussen, B. M., Ghafouri-Fard, S., et al. (2022). Assessment of ACE1 Variants and ACE1/ACE2 Expression in COVID-19 Patients. *Vasc. Pharmacol.* 142, 106934. doi:10.1016/j.vph.2021.106934

Aladag, E., Tas, Z., Ozdemir, B. S., Akbaba, T. H., Akpınar, M. G., Goker, H., et al. (20212021). Human Ace D/I Polymorphism Could Affect the Clinicobiological Course of COVID-19. *J. Renin-Angiotensin-Aldosterone Syst.* 2021, 1–7. doi:10.1155/2021/5509280

Alghamdi, J., Alaamery, M., Barhoumi, T., Rashid, M., Alajmi, H., Aljasser, N., et al. (2021). Interferon-induced Transmembrane Protein-3 Genetic Variant Rs12252 Is Associated with COVID-19 Mortality. *Genomics* 113, 1733–1741. doi:10.1016/j.ygeno.2021.04.002

Amini-Bavil-Olyaee, S., Choi, Y. J., Lee, J. H., Shi, M., Huang, I.-C., Farzan, M., et al. (2013). The Antiviral Effector IFITM3 Disrupts Intracellular Cholesterol Homeostasis to Block Viral Entry. *Cell Host & Microbe* 13, 452–464. doi:10.1016/j.chom.2013.03.006

Annunziata, A., Coppola, A., di Spirito, V., Cauteruccio, R., Marotta, A., Micco, P. D., et al. (2021). The Angiotensin Converting Enzyme Deletion/Deletion Genotype Is a Risk Factor for Severe COVID-19: Implication and Utility for Patients Admitted to Emergency Department. *Medicina* 57, 844. doi:10.3390/medicina57080844

Araújo, J. L., Menezes, D., Saraiva-Duarte, J. M., Ferreira, L., Aguiar, R., and Souza, R. (2021). Systematic Review of Host Genetic Association with Covid-19 Prognosis and Susceptibility: What Have We Learned in 2020? *Rev. Med. Virol.* 32, e2283. doi:10.1002/rmv.2283

Cafiero, C., Rosapepe, F., Palmirotta, R., Re, A., Ottaiano, M. P., Benincasa, G., et al. (2021). Angiotensin System Polymorphisms' in SARS-CoV-2 Positive Patients: Assessment between Symptomatic and Asymptomatic Patients: A Pilot Study. *Pgpm* 14, 621–629. doi:10.2147/PGPM.S303666

Castellon, R., and Hamdi, H. (2007). Demystifying the ACE Polymorphism: From Genetics to Biology. *Cpd* 13, 1191–1198. doi:10.2174/138161207780618902

Cuesta-Llavona, E., Albaiceta, G. M., García-Clemente, M., Duarte-Herrera, I. D., Amado-Rodríguez, L., Hermida-Valverde, T., et al. (2021). Association between the Interferon-Induced Transmembrane Protein 3 Gene (IFITM3) Rs34481144/Rs12252 Haplotypes and COVID-19. *Curr. Res. Virol. Sci.* 2, 100016. doi:10.1016/j.crviro.2021.100016

Davies, N. G., Jarvis, C. I., Jarvis, C. I., Edmunds, W. J., Jewell, N. P., Diaz-Ordaz, K., et al. (2021). Increased Mortality in Community-Tested Cases of SARS-CoV-2 Lineage B.1.1.7. *Nature* 593, 270–274. doi:10.1038/s41586-021-03426-1

Deb, P., Zannat, K. e., Talukder, S., Bhuiyan, A. H., Jilani, M. S. A., and Saif-Ur-Rahman, K. M. (2022). Association of HLA Gene Polymorphism with Susceptibility, Severity, and Mortality of COVID -19: A Systematic Review. *Hla*. [online ahead of print]. doi:10.1111/tan.14560

Delanghe, J. R., Speeckaert, M. M., and de Buyzere, M. L. (2020). The Host's Angiotensin-Converting Enzyme Polymorphism May Explain Epidemiological Findings in COVID-19 Infections. *Clinica Chim. Acta* 505, 192–193. doi:10.1016/j.cca.2020.03.031

Faria, N. R., Mellan, T. A., Whittaker, C., Claro, I. M., Candido, D. D. S., Mishra, S., et al. (2021). Genomics and Epidemiology of the P.1 SARS-CoV-2 Lineage in Manaus, Brazil. *Science* 372, 815–821. doi:10.1126/science.abh2644

Fernández Villalobos, N. V., Ott, J. J., Klett-Tammen, C. J., Bockey, A., Vanella, P., Krause, G., et al. (2021). Effect Modification of the Association between Comorbidities and Severe Course of COVID-19 Disease by Age of Study Participants: a Systematic Review and Meta-Analysis. *Syst. Rev.* 10, 194. doi:10.1186/s13643-021-01732-3

Fishchuk, L., Rossokha, Z., Pokhylko, V., Cherniavska, Y., Tsvirenko, S., Kovtun, S., et al. (2021). Modifying Effects of TNF-α, IL-6 and VDR Genes on the Development Risk and the Course of COVID-19. Pilot Study. *Drug Metab. Personalized Ther*. [online ahead of print]. doi:10.1515/dmpt-2021-0127

Gómez, J., Albaiceta, G. M., Cuesta-Llavona, E., García-Clemente, M., López-Larrea, C., Amado-Rodríguez, L., et al. (2021). The Interferon-Induced Transmembrane Protein 3 Gene (IFITM3) Rs12252 C Variant Is Associated with COVID-19. *Cytokine* 137, 155354. doi:10.1016/j.cyto.2020.155354

Gómez, J., Albaiceta, G. M., García-Clemente, M., López-Larrea, C., Amado-Rodríguez, L., Lopez-Alonso, I., et al. (2020). Angiotensin-converting Enzymes (ACE, ACE2) Gene Variants and COVID-19 Outcome. *Gene* 762, 145102. doi:10.1016/j.gene.2020.145102

Gong, P., Mei, F., Li, R., Wang, Y., Li, W., Pan, K., et al. (2022). Angiotensin-Converting Enzyme Genotype-specific Immune Response Contributes to the Susceptibility of COVID-19: A Nested Case-Control Study. *Front. Pharmacol.* 12, 759587. doi:10.3389/fphar.2021.759587

Gunal, O., Sezer, O., Ustun, G. U., Ozturk, C. E., Sen, A., Yigit, S., et al. (2021). Angiotensin-converting Enzyme-1 Gene Insertion/deletion Polymorphism May Be Associated with COVID-19 Clinical Severity: A Prospective Cohort Study. *Ann. Saudi Med.* 41, 141–146. doi:10.5144/0256-4947.2021.141

Hamdi, H. K., and Castellon, R. (2004). A Genetic Variant of ACE Increases Cell Survival: a New Paradigm for Biology and Disease. *Biochem. Biophysical Res. Commun.* 318, 187–191. doi:10.1016/j.bbrc.2004.04.004

Hashan, M. R., Ghozy, S., El-Qushayri, A. E., Pial, R. H., Hossain, M. A., and al Kibria, G. M. (2021). Association of Dengue Disease Severity and Blood Group: A Systematic Review and Meta-analysis. *Rev. Med. Virol.* 31, 1–9. doi:10.1002/rmv.2147

Hattersley, A. T., and McCarthy, M. I. (2005). What Makes a Good Genetic Association Study? *The Lancet* 366, 1315–1323. doi:10.1016/S0140-6736(05)67531-9

Heidari Nia, M., Rokni, M., Mirinejad, S., Kargar, M., Rahdar, S., Sargazi, S., et al. (2021). Association of Polymorphisms in Tumor Necrosis Factors with SARS-CoV-2 Infection and Mortality Rate: A Case-control Study and In Silico Analyses. *J. Med. Virol* 94, 1502–1512. doi:10.1002/jmv.27477

Hubacek, J. A., Dusek, L., Majek, O., Adamek, V., Cervinkova, T., Dlouha, D., et al. (2021). ACE I/D Polymorphism in Czech First-Wave SARS-CoV-2-Positive Survivors. *Clinica Chim. Acta* 519, 206–209. doi:10.1016/j.cca.2021.04.024

Ioannidis, J. P. A., Ntzani, E. E., Trikalinos, T. A., and Contopoulos-Ioannidis, D. G. (2001). Replication Validity of Genetic Association Studies. *Nat. Genet.* 29, 306–309. doi:10.1038/ng749

Jacobs, M., Lahousse, L., van Eeckhoutte, H. P., Wijnant, S. R. A., Delanghe, J. R., Brusselle, G. G., et al. (2021). Effect of ACE1 Polymorphism Rs1799752 on Protein Levels of ACE2, the SARS-CoV-2 Entry Receptor, in Alveolar Lung Epithelium. *ERJ Open Res.* 7, 00940–02020. doi:10.1183/23120541.00940-2020

Karakaş Çelik, S., Çakmak Genç, G., Pişkin, N., Açikgöz, B., Altinsoy, B., Kurucu İşsiz, B., et al. (2021). Polymorphisms of ACE (I/D) and ACE2 Receptor Gene (Rs2106809, Rs2285666) Are Not Related to the Clinical Course of COVID-19: A Case Study. *J. Med. Virol.* 93, 5947–5952. doi:10.1002/jmv.27160

Karki, R., and Kanneganti, T.-D. (2021). The 'cytokine Storm': Molecular Mechanisms and Therapeutic Prospects. *Trends Immunol.* 42, 681–705. doi:10.1016/j.it.2021.06.001

Konings, F., Perkins, M. D., Kuhn, J. H., Pallen, M. J., Alm, E. J., Archer, B. N., et al. (2021). SARS-CoV-2 Variants of Interest and Concern Naming Scheme Conducive for Global Discourse. *Nat. Microbiol.* 6, 821–823. doi:10.1038/s41564-021-00932-w

Kouhpayeh, H. R., Tabasi, F., Dehvari, M., Naderi, M., Bahari, G., Khalili, T., et al. (2021). Association between Angiotensinogen (AGT), Angiotensin-Converting Enzyme (ACE) and Angiotensin-II Receptor 1 (AGTR1) Polymorphisms and COVID-19 Infection in the Southeast of Iran: a Preliminary Case-Control Study. *Transl Med. Commun.* 6, 26. doi:10.1186/s41231-021-00106-0

Latini, A., Agolini, E., Novelli, A., Borgiani, P., Giannini, R., Gravina, P., et al. (2020). COVID-19 and Genetic Variants of Protein Involved in the SARS-CoV-2 Entry into the Host Cells. *Genes* 11, 1010. doi:10.3390/genes11091010

Li, M., Schifanella, L., and Larsen, P. A. (2021). Alu Retrotransposons and COVID-19 Susceptibility and Morbidity. *Hum. Genomics* 15, 2. doi:10.1186/s40246-020-00299-9

Liu, S., Kong, C., Wu, J., Ying, H., and Zhu, H. (2012). Effect of CCR5-Δ32 Heterozygosity on HIV-1 Susceptibility: A Meta-Analysis. *PLoS ONE* 7, e35020. doi:10.1371/journal.pone.0035020

Mahmood, Z. S., Fadhil, H. Y., Abdul Hussein, T. A., and Ad'hiah, A. H. (2022). Severity of Coronavirus Disease 19: Profile of Inflammatory Markers and ACE (Rs4646994) and ACE2 (Rs2285666) Gene Polymorphisms in Iraqi Patients. *Meta Gene* 31, 101014. doi:10.1016/j.mgene.2022.101014

Mir, M. M., Mir, R., Alghamdi, M. A. A., Alsayed, B. A., Wani, J. I., Alharthi, M. H., et al. (2021). Strong Association of Angiotensin Converting Enzyme-2 Gene Insertion/deletion Polymorphism with Susceptibility to Sars-Cov-2, Hypertension, Coronary Artery Disease and Covid-19 Disease Mortality. *Jpm* 11, 1098. doi:10.3390/jpm11111098

Moher, D., Booth, A., and Stewart, L. (2014). How to Reduce Unnecessary Duplication: Use PROSPERO. *Bjog: Int. J. Obstet. Gy* 121, 784–786. doi:10.1111/1471-0528.12657

Möhlendick, B., Schönfelder, K., Breuckmann, K., Elsner, C., Babel, N., Balfanz, P., et al. (2021). ACE2 Polymorphism and Susceptibility for SARS-CoV-2 Infection and Severity of COVID-19. *Pharmacogenetics and Genomics* 31, 165–171. doi:10.1097/FPC.0000000000000436

Munafò, M. R., Stothart, G., and Flint, J. (2009). Bias in Genetic Association Studies and Impact Factor. *Mol. Psychiatry* 14, 119–120. doi:10.1038/mp.2008.77

Nasserie, T., Hittle, M., and Goodman, S. N. (2021). Assessment of the Frequency and Variety of Persistent Symptoms Among Patients with COVID-19. *JAMA Netw. Open* 4, e2111417. doi:10.1001/jamanetworkopen.2021.11417

Niemi, M. E. K., Karjalainen, J., Liao, R. G., Neale, B. M., Daly, M., Ganna, A., et al. (2021). Mapping the Human Genetic Architecture of COVID-19. *Nature* 600, 472–477. doi:10.1038/s41586-021-03767-x

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 Statement: an Updated Guideline for Reporting Systematic Reviews. *BMJ* 372, n71. doi:10.1136/bmj.n71

Papadopoulou, A., Fragkou, P. C., Maratou, E., Dimopoulou, D., Kominakis, A., Kokkinopoulou, I., et al. (2022). Angiotensin-converting-enzyme Insertion/deletion Polymorphism, ACE Activity, and COVID-19: A rather Controversial Hypothesis. A Case-control Study. *J. Med. Virol.* 94, 1050–1059. doi:10.1002/jmv.27417

Pati, A., Mahto, H., Padhi, S., and Panda, A. K. (2020). ACE Deletion Allele Is Associated with Susceptibility to SARS-CoV-2 Infection and Mortality Rate: An Epidemiological Study in the Asian Population. *Clinica Chim. Acta* 510, 455–458. doi:10.1016/j.cca.2020.08.008

Peacock, T. P., Goldhill, D. H., Zhou, J., Baillon, L., Frise, R., Swann, O. C., et al. (2021). The Furin Cleavage Site in the SARS-CoV-2 Spike Protein Is Required for Transmission in Ferrets. *Nat. Microbiol.* 6, 899–909. doi:10.1038/s41564-021-00908-w

Popejoy, A. B., and Fullerton, S. M. (2016). Genomics Is Failing on Diversity. *Nature* 538, 161–164. doi:10.1038/538161a

Prabhu, S. S., Chakraborty, T. T., Kumar, N., and Banerjee, I. (2018). Association between IFITM3 Rs12252 Polymorphism and Influenza Susceptibility and Severity: A Meta-Analysis. *Gene* 674, 70–79. doi:10.1016/j.gene.2018.06.070

Praissman, J. L., and Wells, L. (2021). Proteomics-Based Insights into the SARS-CoV-2-Mediated COVID-19 Pandemic: A Review of the First Year of Research. *Mol. Cell Proteomics* 20, 100103. doi:10.1016/j.mcpro.2021.100103

Prelli Bozzo, C., Nchioua, R., Volcic, M., Koepke, L., Krüger, J., Schütz, D., et al. (2021). IFITM Proteins Promote SARS-CoV-2 Infection and Are Targets for Virus Inhibition *In Vitro*. *Nat. Commun.* 12, 4584. doi:10.1038/s41467-021-24817-y

R Core Team (2014). *R Core Team (2014). R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: : http://www.R-project.org/.

Rossi, Á. D., de Araújo, J. L. F., de Araújo, J. L. F., de Almeida, T. B., Ribeiro-Alves, M., de Almeida Velozo, C., et al. (2021). Association between ACE2 and TMPRSS2 Nasopharyngeal Expression and COVID-19 Respiratory Distress. *Sci. Rep.* 11, 9658. doi:10.1038/s41598-021-88944-8

Saad, H., Jabotian, K., Sakr, C., Mahfouz, R., Akl, I. B., and Zgheib, N. K. (2021). The Role of Angiotensin Converting Enzyme 1 Insertion/Deletion Genetic Polymorphism in the Risk and Severity of COVID-19 Infection. *Front. Med.* 8, 798571. doi:10.3389/fmed.2021.798571

Sagoo, G. S., Little, J., and Higgins, J. P. T. (2009). Systematic Reviews of Genetic Association Studies. *Plos Med.* 6, e1000028. doi:10.1371/journal.pmed.1000028

Saleh, A., Sultan, A., Elashry, M. a., Farag, A., Mortada, M. I., Ghannam, M. A., et al. (2020). Association of TNF-α G-308 a Promoter Polymorphism with the Course and Outcome of COVID-19 Patients. *Immunological Invest.*, 1–12. doi:10.1080/08820139.2020.1851709

Samprathi, M., and Jayashree, M. (2021). Biomarkers in COVID-19: An Up-To-Date Review. *Front. Pediatr.* 8, 607647. doi:10.3389/fped.2020.607647

Schönfelder, K., Breuckmann, K., Elsner, C., Dittmer, U., Fistera, D., Herbstreit, F., et al. (2021). The Influence of IFITM3 Polymorphisms on Susceptibility to SARS-CoV-2 Infection and Severity of COVID-19. *Cytokine* 142, 155492. doi:10.1016/j.cyto.2021.155492

Sohani, Z. N., Sarma, S., Alyass, A., de Souza, R. J., Robiou-du-Pont, S., Li, A., et al. (2016). Empirical Evaluation of the Q-Genie Tool: a Protocol for Assessment of Effectiveness. *BMJ Open* 6, e010403. doi:10.1136/bmjopen-2015-010403

Spencer, C. C. A., Su, Z., Donnelly, P., and Marchini, J. (2009). Designing Genome-wide Association Studies: Sample Size, Power, Imputation, and the Choice of Genotyping Chip. *Plos Genet.* 5, e1000477. doi:10.1371/journal.pgen.1000477

Sriram, K., and Insel, P. A. (2020). A Hypothesis for Pathobiology and Treatment of COVID-19 : The Centrality of ACE1/ACE2 Imbalance. *Br. J. Pharmacol.* 177, 4825–4844. doi:10.1111/bph.15082

Stebbing, J., Phelan, A., Griffin, I., Tucker, C., Oechsle, O., Smith, D., et al. (2020). COVID-19: Combining Antiviral and Anti-inflammatory Treatments. *Lancet Infect. Dis.* 20, 400–402. doi:10.1016/S1473-3099(20)30132-8

Suehiro, T., Morita, T., Inoue, M., Kumon, Y., Ikeda, Y., and Hashimoto, K. (2004). Increased Amount of the Angiotensin-Converting Enzyme (ACE) mRNA Originating from the ACE Allele with Deletion. *Hum. Genet.* 115, 91. doi:10.1007/s00439-004-1136-4

Sullivan, P. F. (2007). Spurious Genetic Associations. *Biol. Psychiatry* 61, 1121–1126. doi:10.1016/j.biopsych.2006.11.010

Torre-Fuentes, L., Matías-Guiu, J., Hernández-Lorenzo, L., Montero-Escribano, P., Pytel, V., Porta-Etessam, J., et al. (2021). ACE2, TMPRSS2 , and Furin Variants and SARS-CoV-2 Infection in Madrid, Spain. *J. Med. Virol.* 93, 863–869. doi:10.1002/jmv.26319

Verma, S., Abbas, M., Verma, S., Khan, F. H., Raza, S. T., Siddiqi, Z., et al. (2021). Impact of I/D Polymorphism of Angiotensin-Converting Enzyme 1 (ACE1) Gene on the Severity of COVID-19 Patients. *Infect. Genet. Evol.* 91, 104801. doi:10.1016/j.meegid.2021.104801

Yamamoto, N., Ariumi, Y., Nishida, N., Yamamoto, R., Bauer, G., Gojobori, T., et al. (2020). SARS-CoV-2 Infections and COVID-19 Mortalities Strongly Correlate with ACE1 I/D Genotype. *Gene* 758, 144944. doi:10.1016/j.gene.2020.144944

Zhang, Y., Qin, L., Zhao, Y., Zhang, P., Xu, B., Li, K., et al. (2020). Interferon-Induced Transmembrane Protein 3 Genetic Variant Rs12252-C Associated with Disease Severity in Coronavirus Disease 2019. *J. Infect. Dis.* 222, 34–37. doi:10.1093/infdis/jiaa224

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

CAPÍTULO 3

Neste capítulo, conduzimos ensaios de expressão gênica para os genes *ACE1*, *ACE2* e *TMPRSS2,* além da genotipagem do polimorfismo rs4646994 (ALU287pb) no gene *ACE1* em uma amostra de pacientes COVID-19 positivos hospitalizados em Belo Horizonte e Rio de Janeiro. Nosso objetivo foi investigar a influência destes marcadores no desfecho clínico da COVID-19. Também realizamos uma metanálise combinando efeitos das análises com as amostras do Rio de Janeiro e Belo Horizonte para a dificuldade respiratória, considerando apenas indivíduos hospitalizados. Realizamos uma metanálise combinando o efeito observado na análise de associação do polimorfismo rs4646994 e a chance de óbito na amostra de Belo Horizonte com estudos publicados na literatura, e conduzimos também uma metanálise combinando os efeitos observados na análise de associação do polimorfismo rs4646994 na amostra do Rio de Janeiro com a chance de hospitalização com estudos publicados na literatura.

Evaluation of COVID-19 severity and mortality association with *ACE1* Alu 287 bp polymorphism and *ACE1*, *ACE2*, and *TMPRSS2* expression in hospitalised patients

João Locke Ferreira de Araújo[1]; Átila Duque Rossi[2]; Hugo José Alves[1]; Renata Eliane de Ávila[3]; Gustavo Gomes Resende[4]; Mauro Martins Teixeira[5], Renato Santana Aguiar [1,6]; Cynthia Chester Cardoso[2]; Renan Pedra de Souza [1,*]

1 Grupo de Pesquisa em Bioestatística e Epidemiologia Molecular, Laboratório de Biologia Integrativa, Departamento de Genética, Ecologia e Evolução; Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

2 Laboratório de Virologia Molecular, Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil

3 Hospital Eduardo de Menezes, Belo Horizonte, MG, Brazil

4 Hospital das Clínicas, Universidade Federal de Minas Gerais (HC-UFMG/EBSERH), Belo Horizonte, MG, Brazil

5 Departamento de Bioquímica e Imunologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, UFMG, Brazil.

6 Instituto D'OR de Pesquisa e Ensino, Rio de Janeiro, RJ, Brazil

[*] Corresponding author: Renan P. Souza (renanpedra@gmail.com) Universidade Federal de Minas Gerais. Av. Antônio Carlos, 6627 ICB – Pampulha, 31270901 – Belo Horizonte – Minas Gerais – Brazil. Phone: +553134092895.

Abstract

The angiotensin-converting enzyme 2 (ACE2) and the transmembrane serine two protease (TMPRSS2) are central human molecules in the SARS-CoV-2 virus-host interaction. The angiotensin-converting enzyme 1 (ACE1) catalyses the conversion of angiotensin I to angiotensin II, an ACE2 substrate. Evidence indicates that *ACE1* may influence *ACE2* expression. Identifying biomarkers associated with COVID-19 outcomes will help clarify its pathophysiology and improve prognosis. Therefore, we investigated whether *ACE1*, *ACE2,* and *TMPRSS2* gene expression, and the *ACE1* Alu 287 bp polymorphism (rs4646994), contributed to COVID-19 severity and mortality in hospitalised patients. Severity was defined as the need for mechanical ventilation. Samples were collected in two Brazilian cities in 2020: Belo Horizonte (n=134) and Rio de Janeiro (n=41). A sample of mild patients in Rio de Janeiro who were not hospitalised (n=172) was also collected. The median age differed between clinical sites ($p = 0.016$), and no difference in median days of hospitalisation was observed ($p = 0.329$). Age was associated with severity ($p = 0.014$) and mortality ($p = 0.014$) in the Belo Horizonte cohort. Age was associated with the chance of hospitalisation ($p<0.001$) in the Rio de Janeiro cohort. No alteration in *ACE1*, *ACE2* and *TMPRSS2* expression was associated with severity or mortality. *ACE1* polymorphism did not influence the likelihood of both outcomes. There was also no association with the chance of hospitalization, although the meta-analysis involving our original data with the literature showed an effect: I-allele dominance was 0.76 (95% confidence interval: 0.61-0.95) while for D-allele dominance was 1.39 (95 % confidence interval: 1.12-1.72). Additional investigations assessing multiple candidate genes are crucial to understanding molecular mechanisms influencing COVID-19 prognosis due to its multifactorial structure.

Keywords: molecular epidemiology; indel; genetic association; genetic variability; biomarkers

Introduction

Coronavirus 2019 disease (COVID-19) is caused by a virus of the coronaviridae family, known as the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). COVID-19 clinical manifestation can be highly heterogenous with patients ranging from asymptomatic to severe cases. Several clinical and epidemiological factors seem to contribute to COVID-19 severity in Brazil [1–3]. From a molecular perspective, genetic differences also seem to contribute to COVID-19 outcomes [1, 4].

The angiotensin-converting enzyme 2 (ACE2) and the transmembrane serine two protease (TMPRSS2) are central human molecules in the virus-host interaction [5]. The spike viral protein interacts with the ACE2 receptor, and the TMPRSS2 cleaves the spike protein's receptor binding domain (RBD), exposing the fusion peptide [6]. Preliminary studies have explored the association between *ACE2* and *TMPRSS2* gene expression and their polymorphisms with COVID-19 outcomes[7–10]. Significant expression alterations were found in subjects presenting respiratory distress [7].

The angiotensin-converting enzyme 1 (ACE1) catalyses the conversion of angiotensin I to angiotensin II, an ACE2 substrate. Evidence indicates that *ACE1* may influence *ACE2* expression[11]. An *ACE1* 287bp insertion/deletion polymorphism (rs4646994) has been associated with increased *ACE1* enzyme activity in individuals with a genotype homozygous for the deletion (D/D) [12]. A recent metanalysis showed a 45% increase in the chance of severe COVID-19 manifestation in deletion carriers (D allele), although no effect on susceptibility was found [1].

Identifying biomarkers associated with COVID-19 outcomes will help clarify its pathophysiology and improve prognosis. Proteins related to virus-host interaction are strong candidates for biomarkers. Therefore, we evaluated whether *ACE1*, *ACE2,* and *TMPRSS2* gene expression and *ACE1* polymorphism (Alu 287 bp) would contribute to the severity in hospitalised COVID-19 patients in a Brazilian sample.

Methods

Enrolled subjects were inpatients from two Brazilian hospitals: Hospital Naval Marcilio Dias in Rio de Janeiro (n = 41) and Eduardo de Menezes in Belo Horizonte (n = 134). Furthermore, 172 patients with mild symptoms collected at Universidade Federal do Rio de Janeiro were included in a second cohort in Rio de Janeiro for genetic association study with alu 287bp. All analysed biological material was obtained from positive residual diagnostic samples in 2020 before reporting the SARS-CoV-2 variants of concern. Rio de Janeiro samples were nasopharyngeal swabs, while Belo Horizonte samples were nasopharyngeal swabs (n = 102) and bronchoalveolar lavage (n = 32). The study was conducted according to the Declaration of Helsinki and approved by the Ethics Committee (protocols 30161620.0.0000.5257, 32382820.3.0000.5256, 32224420.3.0000.0008, 31462820.3.0000.5149). We explored biomarker effects in two outcomes: the need for mechanical ventilation during hospitalisation (using both samples) and mortality (using the Belo Horizonte sample since no deaths were recorded in the Rio de Janeiro sample). All molecular experiments were conducted blinded to the outcome information.

Samples were collected in a viral transport medium and stored at -80ºC until extraction. RNA and DNA extractions were performed with a Quick-RNA Viral kit (Zymo Research, CA, USA), following the manufacturer's instructions and standardised laboratory protocol. cDNA was synthesised using the high-capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, MA, USA) according to the manufacturer's instructions.

ACE1, *ACE2*, *TMPRSS2,* and *B2M* (reference gene) expression levels were evaluated in 40 cycles of quantitative PCR with Integrated DNA Technologies (NJ, USA) exon-exon junction probes (Hs.PT.58.19167084, Hs.PT.58.27645939, HS.PT.58.39738666 and Hs.PT.39a.22214847). The ΔCT was calculated by subtracting the cycle threshold (CT) of the gene of interest from *B2M* CT. All samples that amplified the reference gene were included in the analysis. Samples without amplification of the target gene were analysed considering CT equals 40 (minimum expression level). Gene expression analysis included only samples from nasopharyngeal swabs.

ACE1 Alu 287 bp polymorphism was genotyped using the FastStart Universal SYBR Green Master kit (Promega, WI, USA) with three primers: 5'CATCCTTTCTCCCATTTCTC3' (Primer1, Forward); 5'TGGGATTACAGGCGTGATACAG3' (Primer 2, Forward, internal); and 5'ATTTCAGAGCTGGAATAAAATT 3' (Primer 3, Reverse) [13]. Primer stock was resuspended at 100µM and diluted a working solution to 10µM. The concentration of primers

1 and 3 in the reaction is 20 picomol, and the concentration of primer 2 is 40 picomol. The generated fragment sizes were 65bp (Insertion) and 84bp (deletion). The visualisation was performed in 3% agarose gel. 10% of the sample was randomly genotyped to assess the genotyping's quality. The agreement's level was 100%.

Statistical analyses were performed using the R program (version 4.1.2). Data from clinical sites were compared using the Wilcoxon and Fisher's Exact tests. Hardy-Weinberg equilibrium deviations were evaluated separately in cases and controls using Pearson's chi-squared test in the *SNPassoc* package. Median delta Ct differences were evaluated using the Wilcoxon test. Genetic association with outcomes was assessed with Pearson's chi-squared or Fisher's Exact tests, respecting test assumptions. Combined polymorphism effects were obtained with metanalysis using the Mantel–Haenszel weighted means method under the fixed-effect model implemented on the *metabin* function (*meta* package). A significance level was set at 5%.

Results

Collected data were compared between clinical sites. A difference in median age was observed ($p = 0.016$) with no difference in median days of hospitalisation ($p = 0.329$) (Figure 1). Most evaluated symptoms were homogeneously distributed, except for adynamia and vomit (Table 1). Clinical outcomes also showed significance between sites, with the Belo Horizonte site presenting increased severity, as shown by the association of admission to the intensive care unit, respiratory support type and death (Table 1).

Median *ACE1*, *ACE2,* and *TMPRSS2* gene expression did not significantly differ according to both investigated outcomes (the need for mechanical ventilation and death) in hospitalised patients (Table 2). Furthermore, the median ratio between *ACE2* and *TMPRSS2* expression did not show an effect. As expected, increased median age was found among subjects who died compared to those who survived.

No Hardy-Weinberg equilibrium violations were observed ($p > 0.05$ for all samples, data not shown). No association was found between *ACE1* Alu 287 bp polymorphism and both outcomes (Table 3). When testing hospitalized versus non-hospitalized patients in the Rio de Janeiro sample, there was a difference in age ($p < 0.001$) although no association was observed for Alu 287 bp polymorphism either (Table 4). Combined effects from both samples on the need for mechanical ventilation also did not reach significance: pooled odds-ratio for D-allele dominance was 1.15 (95% confidence interval: 0.59-2.25) while for I-allele

dominance was 0.96 (95% confidence interval: 0.43-2.10). Since the number of subjects varied from the expression analysis, we reevaluated the age effect and observed a significant median difference in the Belo Horizonte sample for both outcomes.

We carried out a literature search in the Pubmed database, complementary to the work already carried out by Araújo et al [1] with the aim of observing the combined effects of the Belo Horizonte sample on mortality and for the Rio de Janeiro sample on the need for hospitalisation with other studies published in the literature. The search was performed on November 10, 2022. For mortality, the meta-analysis was performed with two more studies[14, 15], in which we also did not observe significance: pooled odds-ratio for I-allele dominance was 0.63 (95% confidence interval: 0.20-1.93) while for D-allele dominance was 1.48 (95% confidence interval: 0.38-5.81).

We also checked the combined effect of the Rio de Janeiro sample comparing mild patients with those who required hospitalisation with the literature. The meta-analysis was conducted with seven more studies extracted from the literature [16–22], where we observed significance: pooled odds-ratio for I-allele dominance was 0.76 (95% confidence interval: 0.61-0.95) while for D-allele dominance was 1.39 (95 % confidence interval: 1.12-1.72).

Discussion

Molecular signatures associated with COVID-19-related outcomes have been extensively investigated during the pandemic. Molecules related to the immune response have been, by far, the most studied. Among the most significant results, an association was reported between circulating interleukin-6 and COVID-19 severity in a metanalysis combining 15 original studies[23]. Proteins associated with virus-host interaction can also be promissing candidates for biomarker studies.

ACE2 and *TMPRSS2* expressions have been explored due to their central role in the cell entry mechanisms. Higher ACE2 protein levels were found in post-mortem lung samples of patients who died of severe COVID-19 suggesting a pathobiological role in disease severity[24]. *TMPRSS2*/*ACE2* expression ratio was associated with respiratory distress [7]. Age-dependent *ACE2* gene expression in the nasal epithelium could explain why children have lower infection susceptibility and mortality than adults [25]. However, a recent study did not find differences between infants and adults assessing ACE2 immunofluorescence staining and protein levels[26]. We report no significant association between *ACE2* and *TMPRSS2* gene expression and the need for mechanical ventilation or death. Similarly, no

ACE2 expression differences were found between those admitted to the intensive care unit and patients who were not [27].

ACE1 also seems a good biomarker candidate, although not directly related to viral cell entry. ACE1/ACE2 balance has been hypothesised to contribute to clinical phenotypes relevant to COVID-19 [28, 29]. ACE1 inhibitors were associated with a significantly reduced risk of hospital admission during COVID-19 in a cohort study including 8.3 million people[30]. We did not find altered *ACE1* expression, although a previous study reported *ACE1* expression was significantly higher in COVID-19 intensive care unit patients[4]. Similarly, no association between *ACE1* Alu 287 bp polymorphism and COVID-19 severity was achieved.

Our report presents limitations. First, replications are warranted because the study may be underpowered to detect minor effects. Second, we could not evaluate the viral diversity impact since samples were collected before describing the variants of concern that substantially changed COVID-19 severity[31]. Another relevant factor that could not be unexplored was the vaccination status. Therefore, additional investigations in larger samples from diverse ethnic backgrounds assessing multiple candidate genes are crucial to understanding COVID-19 prognosis due to its multifactorial structure.

Conflict of interest

The authors did not show any conflict of interest.

Funding

We acknowledge support from the Rede Corona-ômica BR MCTI/FINEP affiliated with RedeVírus/MCTI (01.20.0029.000462/20 404096/2020-4; 1227/21 01.22.0074.00); Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (315592/2021-4); Financiadora de Estudos e Projetos - FINEP (0494/20 01.20.0026.00; 1228/21 01.22.0082.00; 1139/20 01.20.0076.00); Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES (Finance Code 001).

Reference list

1. de Araújo JLF, Menezes D, de Aguiar RS, de Souza RP (2022) IFITM3, FURIN, ACE1, and TNF-α Genetic Association With COVID-19 Outcomes: Systematic Review and Meta-Analysis. Front. Genet. 13:775246

2. Brizzi A, Whittaker C, Servo LMS, et al (2022) Spatial and temporal fluctuations in COVID-19 fatality rates in Brazilian hospitals. Nat Med. https://doi.org/10.1038/s41591-022-01807-1

3. Marcolino MS, Ziegelmann PK, Souza-Silva MVR, et al (2021) Clinical characteristics and outcomes of patients hospitalized with COVID-19 in Brazil: results from the Brazilian COVID-19 Registry. Int J Infect Dis. https://doi.org/10.1016/j.ijid.2021.01.019

4. Akbari M, Taheri M, Mehrpoor G, et al (2022) Assessment of ACE1 variants and ACE1/ACE2 expression in COVID-19 patients. Vascul Pharmacol 142:106934. https://doi.org/10.1016/j.vph.2021.106934

5. Muus C, Luecken MD, Eraslan G, et al (2021) Single-cell meta-analysis of SARS-CoV-2 entry genes across tissues and demographics

6. Hoffmann M, Kleine-Weber H, Schroeder S, et al (2020) SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 181:271-280.e8. https://doi.org/10.1016/j.cell.2020.02.052

7. Rossi ÁD, de Araújo JLF, de Almeida TB, et al (2021) Association between ACE2 and TMPRSS2 nasopharyngeal expression and COVID-19 respiratory distress. Sci Rep 11:9658. https://doi.org/10.1038/s41598-021-88944-8

8. Initiative HG, Covid- T, Initiative HG, et al (2022) Matters arising A first update on mapping the human genetic architecture of COVID-19. 608:

9. Saengsiwaritt W, Jittikoon J, Chaikledkaew U, Udomsinprasert W (2022) Genetic polymorphisms of ACE1, ACE2, and TMPRSS2 associated with COVID-19 severity: A systematic review with meta-analysis. Rev Med Virol 32:e2323. https://doi.org/10.1002/rmv.2323

10. Taglauer ES, Wachman EM, Juttukonda L, et al (2022) Acute Severe Acute Respiratory Syndrome Coronavirus 2 Infection in Pregnancy Is Associated with Placental Angiotensin-Converting Enzyme 2 Shedding. Am J Pathol 192:595–603. https://doi.org/10.1016/j.ajpath.2021.12.011

11. Hamdi HK, Castellon R (2004) A genetic variant of ACE increases cell survival: a new paradigm for biology and disease. Biochem Biophys Res Commun 318:187–191. https://doi.org/10.1016/j.bbrc.2004.04.004

12. Suehiro T, Morita T, Inoue M, et al (2004) Increased amount of the angiotensin-converting enzyme (ACE) mRNA originating from the ACE allele with deletion. Hum Genet 115:91–96. https://doi.org/10.1007/s00439-004-1136-4

13. Evans AE, Poirier O, Kee F, et al (1994) Polymorphisms of the angiotensin-converting-enzyme gene in subjects who die from coronary heart disease. Q J Med 87:211–214

14. Mir MM, Mir R, Alghamdi MAA, et al (2021) Strong association of angiotensin converting enzyme-2 gene insertion/deletion polymorphism with susceptibility to sars-cov-2, hypertension, coronary artery disease and COVID-19 disease mortality. J Pers Med 11:1–21. https://doi.org/10.3390/jpm11111098

15. Möhlendick B, Schönfelder K, Breuckmann K, et al (2021) ACE2 polymorphism and susceptibility for SARS-CoV-2 infection and severity of COVID-19. Pharmacogenet Genomics 165–171. https://doi.org/10.1097/FPC.0000000000000436

16. Gunal O, Sezer O, Ustun GU, et al (2021) Angiotensin-converting enzyme-1 gene insertion/deletion polymorphism may be associated with COVID-19 clinical severity: a prospective cohort study. Ann Saudi Med 41:141–146. https://doi.org/10.5144/0256-4947.2021.141

17. Kouhpayeh HR, Tabasi F, Dehvari M, et al (2021) Association between angiotensinogen (AGT), angiotensin-converting enzyme (ACE) and angiotensin-II receptor 1 (AGTR1) polymorphisms and COVID-19 infection in the southeast of Iran: a preliminary case-control study. Transl Med Commun 6:. https://doi.org/10.1186/s41231-021-00106-0

18. Mahmood ZS, Fadhil HY, Abdul Hussein TA, Ad'hiah AH (2022) Severity of coronavirus disease 19: Profile of inflammatory markers and ACE (rs4646994) and ACE2 (rs2285666) gene polymorphisms in Iraqi patients. Meta gene 31:101014. https://doi.org/10.1016/j.mgene.2022.101014

19. Saad H, Jabotian K, Sakr C, et al (2021) The Role of Angiotensin Converting Enzyme 1 Insertion/Deletion Genetic Polymorphism in the Risk and Severity of COVID-19 Infection. Front Med 8:1–13. https://doi.org/10.3389/fmed.2021.798571

20. Verma S, Abbas M, Verma S, et al (2021) Impact of I/D polymorphism of angiotensin-converting enzyme 1 (ACE1) gene on the severity of COVID-19 patients. Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis 91:104801. https://doi.org/10.1016/j.meegid.2021.104801

21. Gong P, Mei F, Li R, et al (2022) Angiotensin-Converting Enzyme Genotype–Specific Immune Response Contributes to the Susceptibility of COVID-19: A Nested Case–Control Study. Front Pharmacol 12:1–13. https://doi.org/10.3389/fphar.2021.759587

22. Martínez-Gómez LE, Herrera-López B, Martinez-Armenta C, et al (2022) ACE and ACE2 Gene Variants Are Associated With Severe Outcomes of COVID-19 in Men. Front Immunol 13:1–10. https://doi.org/10.3389/fimmu.2022.812940

23. Zawawi A, Naser AY, Alwafi H, Minshawi F (2021) Profile of Circulatory Cytokines and Chemokines in Human Coronaviruses: A Systematic Review and Meta-Analysis. Front Immunol 12:666223. https://doi.org/10.3389/fimmu.2021.666223

24. Gheware A, Ray A, Rana D, et al (2022) ACE2 protein expression in lung tissues of severe COVID-19 infection. Sci Rep 12:4058. https://doi.org/10.1038/s41598-022-07918-6

25. Bunyavanich S, Do A, Vicencio A (2020) Nasal Gene Expression of Angiotensin-Converting Enzyme 2 in Children and Adults. JAMA 323:2427–2429. https://doi.org/10.1001/jama.2020.8707

26. Zhu Y, Chew KY, Wu M, et al (2022) Ancestral SARS-CoV-2, but not Omicron, replicates less efficiently in primary pediatric nasal epithelial cells. PLoS Biol 20:e3001728. https://doi.org/10.1371/journal.pbio.3001728

27. Akbari H, Tabrizi R, Lankarani KB, et al (2020) The role of cytokine profile and lymphocyte subsets in the severity of coronavirus disease 2019 (COVID-19): A systematic review and meta-analysis. Life Sci 258:118167. https://doi.org/10.1016/j.lfs.2020.118167

28. Brosnihan KB, Neves LAA, Chappell MC (2005) Does the angiotensin-converting enzyme (ACE)/ACE2 balance contribute to the fate of angiotensin peptides in programmed hypertension? Hypertens. (Dallas, Tex. 1979) 46:1097–1099

29. Mizuiri S, Hemmi H, Arita M, et al (2008) Expression of ACE and ACE2 in individuals with diabetic kidney disease and healthy controls. Am J kidney Dis Off J Natl Kidney Found 51:613–623. https://doi.org/10.1053/j.ajkd.2007.11.022

30. Hippisley-Cox J, Young D, Coupland C, et al (2020) Risk of severe COVID-19 disease with ACE inhibitors and angiotensin receptor blockers: cohort study including 8.3 million people. Heart 106:1503–1511. https://doi.org/10.1136/heartjnl-2020-317393

31. Telenti A, Hodcroft EB, Robertson DL (2022) The Evolution and Biology of SARS-CoV-2 Variants. Cold Spring Harb Perspect Med 12:. https://doi.org/10.1101/cshperspect.a041390

Table 1: Comparison of clinical and epidemiological data between clinical sites. Data is presented with absolute and relative frequencies. n: sample size. BAL - brochoalveolar lavage. [a] Rio de Janeiro (swab) x Belo Horizonte (swab + Bal): samples used in the *ACE1* polymorphism analysis. [b] Rio de Janeiro (swab) x Belo Horizonte (swab): samples used in the gene expression analysis.

Variable	Rio de Janeiro (swab), n = 41	Belo Horizonte (swab+BAL), n = 134	p-value[a]	Belo Horizonte (swab only), n = 102	p-value[b]
Sample from swab - n (%)	41 (100%)	102 (76%)	-	102 (100%)	0.999
Female - n (%)	26 (63%)	66 (49%)	0.112	54 (53%)	0.254
Comorbidity - n (%)	27 (66%)	93 (69%)	0.668	69 (68%)	0.836
Chronic medication use - n (%)	23 (79%)	95 (71%)	0.358	69 (68%)	0.225
Fever - n (%)	33 (80%)	104 (78%)	0.754	80 (79%)	0.864
Chills - n (%)	2 (4.9%)	5 (3.7%)	0.667	4 (3.9%)	0.999
Cough - n (%)	31 (76%)	111 (83%)	0.301	87 (85%)	0.168
Sneezing - n (%)	5 (12%)	15 (12%)	0.999	9 (9.2%)	0.554
Dyspnea - n (%)	34 (83%)	114 (85%)	0.739	84 (82%)	0.935
Coryza - n (%)	7 (17%)	42 (31%)	0.075	32 (31%)	0.083
Headache - n (%)	11 (27%)	42 (31%)	0.582	36 (35%)	0.330
Adynamia - n (%)	4 (9.8%)	89 (66%)	**<0.001**	58 (57%)	**<0.001**
Nausea - n (%)	4 (9.8%)	18 (13%)	0.534	16 (16%)	0.355
Vomit - n (%)	2 (4.9%)	24 (18%)	**0.039**	19 (19%)	**0.034**
Diarrhea - n (%)	7 (17%)	33 (25%)	0.313	25 (25%)	0.335
Myalgia - n (%)	19 (46%)	54 (40%)	0.492	50 (49%)	0.772
Anosmia - n (%)	7 (17%)	20 (15%)	0.739	18 (18%)	0.935
Ageusia - n (%)	5 (12%)	11 (8.2%)	0.535	10 (9.8%)	0.764
Fatigue - n (%)	11 (27%)	27 (20%)	0.364	18 (18%)	0.217
Intensive care unit - n (%)	10 (26%)	76 (57%)	**<0.001**	48 (47%)	**0.027**
Respiratory support - any - n (%)	38 (93%)	133 (99%)	**0.041**	101 (99%)	0.071
Respiratory support - catheter - n (%)	22 (54%)	109 (81%)	**<0.001**	92 (90%)	**<0.001**
Respiratory support - mask - n (%)	7 (17%)	63 (47%)	**<0.001**	41 (40%)	**0.008**
Respiratory support - mechanical ventilation - n (%)	9 (22%)	59 (48%)	**0.004**	29 (32%)	0.259
Death - n (%)	0 (0%)	34 (25%)	**<0.001**	14 (14%)	**0.011**

Table 2: *ACE1*, *ACE2* and *TMPRSS2* expression according to patients need for mechanical ventilation or death. n: sample size.

	Need for mechanical ventilation (Rio de Janeiro)			Need for mechanical ventilation (Belo Horizonte)			Death (Belo Horizonte)		
Variable	No, n = 32	Yes, n = 9	p-value	No, n = 63	Yes, n = 29	p-value	No, n = 88	Yes, n = 14	p-value
Age - median (interquartile range) missing data	41 (40, 59) 2	55 (52, 63)	0.291	54 (44, 65) 0	54 (48, 68) 0	0.215	52 (44, 63) 0	68 (64, 82) 0	**<0.001**
ACE1 delta Ct - median (interquartile range) missing data	Not available	Not available	Not available	11.3 (8.4, 13.4) 12	10.3 (7.7, 13.0) 8	0.552	11.3 (8.5, 13.4) 17	8.8 (6.8, 11.1) 3	0.226
ACE2 delta Ct - median (interquartile range) missing data	6.40 (4.97, 7.83) 0	8.65 (5.36, 8.72) 0	0.128	15.1 (12.2, 17.7) 11	12.9 (10.8, 15.9) 3	0.192	13.5 (10.7, 17.5) 13	14.4 (12.4, 15.4) 1	0.888
TMPRSS2 delta Ct - median (interquartile range) missing data	4.57 (3.72, 5.65) 0	4.87 (4.18, 9.38) 0	0.206	9.0 (5.3, 11.8) 11	8.4 (4.5, 13.8) 3	0.845	8.4 (4.9, 12.1) 13	8.4 (5.8, 10.3) 1	0.925
ACE2/TMPRSS2 delta Ct ratio - median (interquartile range) missing data	1.26 (1.16, 1.64)	1.37 (1.19, 1.51)	0.938	1.58 (1.00, 2.33) 11	1.40 (1.00, 2.46) 3	0.582	.54 (1.00, 2.58) 13	1.43 (1.00, 1.92) 1	0.972

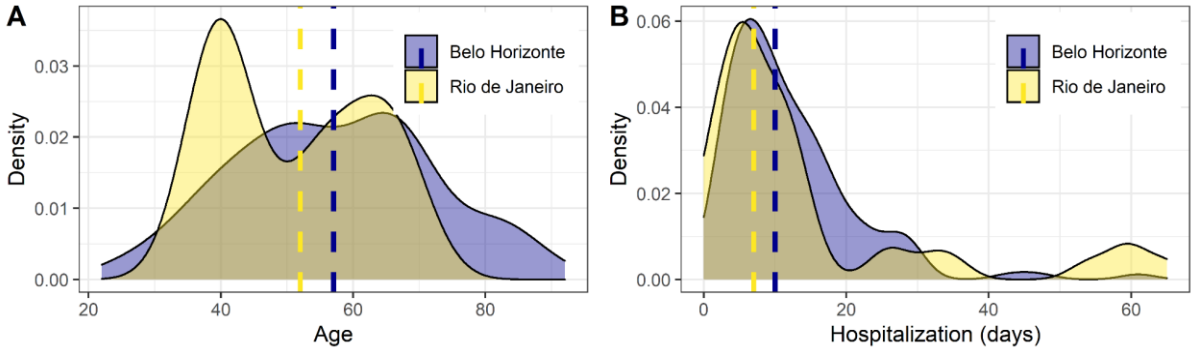
Table 3: Association of *ACE1* Alu 287 bp and patients need for mechanical ventilation or death. Patients whose genotyping reactions did not work explain the differences between sample size and genotype counts. n: sample size.

Variable		Need for mechanical ventilation (Rio de Janeiro)			Need for mechanical ventilation (Belo Horizonte)			Death (Belo Horizonte)		
		No, n = 32	Yes, n = 9	p-value	No, n = 65	Yes, n = 59	p-value	No, n = 100	Yes, n = 34	p-value
Age - median (interquatile range)		41 (40, 59)	55 (52, 63)	0.291	54 (44, 65)	63 (48, 69)	**0.014**	54 (44, 65)	67 (59, 80)	**<0.001**
Co-dominance	D/D - n (%)	10 (31%)	2 (22%)	0.698	23 (36%)	24 (42%)	0.739	39 (39%)	16 (50%)	0.566
	D/I - n (%)	16 (50%)	4 (44%)		27 (42%)	23 (40%)		40 (40%)	11 (34%)	
	I/I - n (%)	6 (19%)	3 (33%)		14 (22%)	10 (18%)		20 (20%)	5 (16%)	
I-allele dominance	DD - n (%)	10 (31%)	2 (22%)	0.702	23 (36%)	24 (42%)	0.487	39 (39%)	16 (50%)	0.291
	II + DI - n (%)	22 (69%)	7 (78%)		41 (64%)	33 (58%)		60 (61%)	16 (50%)	
D-allele dominance	DD + DI - n (%)	26 (81%)	6 (67%)	0.384	50 (78%)	47 (82%)	0.551	79 (80%)	27 (84%)	0.567
	I - n (%)	6 (19%)	3 (33%)		14 (22%)	10 (18%)		20 (20%)	5 (16%)	

Table 4: Association of *ACE1* Alu 287 bp and hospitalized patients on Rio de Janeiro sample. Patients whose genotyping reactions did not work explain the differences between sample size and genotype counts. n: sample size.

Variable		Non-Hospitalized n = 172	Hospitalized n = 41	p- value
Age - median (interquatile range)		39 (30,44)	52 (40,62)	**<0.001**
Co-dominance	D/D - n (%)	54 (31.8%)	12 (29.3%)	
	D/I - n (%)	84 (49.4%)	20 (48.8%)	
	I/I - n (%)	32 (18.8%)	9 (21.9%)	0.891
I-allele dominance	DD - n (%)	54 (31.8%)	12 (29.3%)	
	II + DI - n (%)	116 (68.2%)	29 (70.7%)	0.756
D-allele dominance	DD + DI - n (%)	138 (81.2%)	32 (77.1%)	
	I - n (%)	32 (18.8%)	9 (21.9%)	0.653

Figure 1: Age and hospitalisation days distribution across Belo Horizonte and Rio de Janeiro samples. Dashed lines indicate medians. Median age difference was significant (p = 0.016). No difference was found between median days of hospitalisation (p = 0.329)



CAPÍTULO 4

Neste capítulo replicamos 3 polimorfismos em 2 genes presentes no lócus 3p21.31, que tem apresentado forte associação com o prognóstico da COVID-19 em estudos de escala genômica : *CXCR6* e *LZTFL1*, e analisamos sua influência na severidade da COVID-19, onde consideramos como severidade os dias de hospitalização e a necessidade de suporte respiratório, além de analisarmos também a associação com o óbito. O estudo gerou um Brief Report que apresentamos a seguir.

Replication of 3p21.31 locus (*CXCR6* and *LZTFL1*) association with COVID-19 outcomes in a Brazilian population

João Locke Ferreira de Araújo[1]; Victória Frigério Bonifácio[2]; Lorena Medeiros Batista[2], Renata Eliane de Ávila[3], Renato Santana Aguiar [1,4], Luciana Bastos-Rodrigues[2,*]; Renan Pedra de Souza [1,*]

1 Grupo de Pesquisa em Bioestatística e Epidemiologia Molecular, Laboratório de Biologia Integrativa, Departamento de Genética, Ecologia e Evolução; Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

2 Departamento de Nutrição, Escola de Enfermagem, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

3 Hospital Eduardo de Menezes, Belo Horizonte, MG, Brazil

4 Instituto D'OR de Pesquisa e Ensino, Rio de Janeiro, RJ, Brazil

[*] Corresponding author: Renan P. Souza (renanpedra@gmail.com) or Luciana Bastos-Rodrigues (lu.bastosr@gmail.com). Universidade Federal de Minas Gerais. Av. Antônio Carlos, 6627 ICB – Pampulha, 31270901 – Belo Horizonte – Minas Gerais – Brazil. Phone: +553134092895.

Abstract

The 3p21.31 locus was associated with a severe COVID-19 prognosis. We aimed to replicate this finding using 102 Brazilian hospitalized patients. *LZTFL1* rs10490770, *CXCR6* rs2234355, and *CXCR6* rs2234358 were genotyped. Primary outcomes were the need for mechanical ventilation, hospitalization days, and death. No genetic association was found with the need for mechanical ventilation and hospitalization days. The *CXCR6* rs2234355 was associated with mortality in a codominance model, with A/A genotype carriers presenting an increased chance of death than A/G (OR: 10.5; 95% CI: 1.55- 70.76). Future studies are needed in larger samples and exploring viral diversity and immunization effects.

Keywords: molecular epidemiology; genetic association; genetic variability; biomarkers; SNP; polymorphism.

COVID-19 is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The highly heterogeneous clinical manifestation [1] presents a multifactorial nature. Therefore, genetic factors may contribute to the risk of severe outcomes [2]. Sex, obesity, cardiovascular problems, low economic status, non-European ancestry and blood type have been associated with increased risks of hospitalization [3–8]. SARS-CoV-2 variants were also associated with greater susceptibility, transmission, more severe prognosis and death [9].

The first published genome-wide association study (GWAS) identified the 3p21.31 locus as significantly associated with severe COVID-19 and respiratory failure [10]. This genomic segment was inherited from Neanderthals [11] and contained several genes, including *CXCR6* (C-X-C motif chemokine receptor 6) and *LZTFL1* (leucine zipper transcription factor-like 1). Other GWAS replicated the 3p21.31 locus [8,12,13], with new risk loci being reported (e.g. 9p13.3 and 19q13.12) [14].

CXCR6 encodes a chemokine receptor preferentially expressed in memory T cells [15]. It serves as a cellular gateway for HIV, having polymorphisms associated with the infection described in the literature [16–19]. *LZTFL1* encodes a protein expressed ubiquitously in the cytoplasm, interacting with numerous other proteins. Due to interaction with E-cadherin and the actin cytoskeleton, *LZTFL1* may also function as a tumour suppressor, thereby regulating the transition from epithelial cells to mesenchymal cells [20,21]. *LZTFL1* further replication in candidate-gene studies has been conducted in samples from the UK, Latvia and Colombia [22–24]. Here, we conducted an association study of *CXCR6* and *LZTFL1* variants with respiratory distress and death in hospitalized COVID-19 patients.

We enrolled 102 hospitalized COVID-19 patients at the Hospital Eduardo de Menezes (Belo Horizonte, Brazil) between August and October 2020. The COVID-19 diagnosis was conducted using RT-qPCR in nasopharyngeal swab samples. The UFMG Human Research Ethics Committee approved the project (protocol number 31095820.4.0000.5149). Four mL of whole peripheral blood was collected in an EDTA tube and stored at room temperature (25ºC) until processing. DNA was extracted from leukocyte cells according to the manufacturer’s protocol using the Macherey-Nagel NucleoSpin Blood Kit. DNA

samples were stored at 4ºC. Genotyping for *CXCR6* (rs2234355 and rs2234358) and *LZTFL1* (rs10490770) were carried out according to the manufacturer's protocol using TaqMan probes (Thermo Fisher Scientific) in the CFX Opus 96 (Bio-Rad). Genotyping was performed blinded from clinical data and was repeated in 10% of the samples as a quality control procedure, with a 100% agreement rate.

All statistical analyzes were performed in R version 4.0.2. The significance level was set at 0.05. Medians were compared using Wilcoxon or Kruskal-Wallis tests, while associations were evaluated with Pearson's Chi-squared and Fisher's exact test (*gtsummary* package). The Hardy-Weinberg equilibrium (HWE) was assessed using Fisher's exact test (*SNPassoc* package). Genetic associations under different dominance models were generated for mortality and the need for mechanical ventilation (*SNPassoc* package).

Patients had a median age of 60 (Table 1). Most patients were male (n = 60, 55%). No significant differences or associations were observed for both outcomes (the need for mechanical ventilation and death) across clinical-epidemiological variables, apart from self-reported ethnicity association with death (p=0.015).

HWE deviation for the *CXCR6* rs2234355 variant in subjects who died was identified (p=0.001) (Table 2). No genetic association was found with the need for mechanical ventilation. A significant association was observed for the *CXCR6* rs2234355 with mortality (p=0.022) in a codominance model (Table 2). A/A genotype carriers presented an increased chance of death than A/G (OR: 10.5; 95% CI: 1.55- 70.76 p = 0.016). No deviations from the HWE were observed (rs2234355 p=0.059; rs2234358 p=0.156; rs10490770 p=0.595) in the overall sample. No significant association was found between genetic variants and length of hospitalization (Table 3).

Multiple factors are involved in the COVID-19 clinical manifestation. Genomic-scale studies have raised promising candidate loci for association studies, and the replication of these loci in distinct and representative samples from different populations is essential to understand the actual contribution of these biomarkers to the COVID-19 prognosis.

After the initial description [10,12], replications of the 3p21.31 region were successful. *LZTFL1* rs11385942 was associated with hospitalization risk in a

Colombian sample [24]. *LZTFL1* loci (rs11385942, rs71325088 and rs73064425) altered COVID-19 susceptibility odds in Latvians [23]. No replications were attempted with the *CXCR6* variants. However, lower *CXCR6* expression was reported in lung tissue cells in patients with severe COVID-19 [25]. One GWAS in a Brazilian population failed to find a hit in the 3p21.31 region [26].

Ancestry is a crucial factor in genetic association studies. While molecular-estimated ethnicity is more accurate [27, 28], we relied on self-reported ethnicity. Nevertheless, the overrepresentation of black and brown reports corroborate official statistics for the Brazilian population. Future studies with larger samples and including a sample ancestry analysis are needed to understand the contribution of genetic biomarkers in clinical prognosis. In addition, it is necessary to replicate these studies in vaccinated subjects and those infected by SARS-CoV-2 variants of concern since viral variability seems to contribute to prognosis [9,30].

Conflict of interest

The authors did not show any conflict of interest.

Funding

We acknowledge support from the Rede Corona-ômica BR MCTI/FINEP affiliated with RedeVírus/MCTI (01.20.0029.000462/20 404096/2020-4; 1227/21 01.22.0074.00); Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (315592/2021-4); Financiadora de Estudos e Projetos - FINEP (0494/20 01.20.0026.00; 1228/21 01.22.0082.00; 1139/20 01.20.0076.00); Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES (Finance Code 001) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais—FAPEMIG (APQ-00475-20).

References

1. Oran, D.P.; Topol, E.J. The Proportion of SARS-CoV-2 Infections That Are Asymptomatic : A Systematic Review. *Ann. Intern. Med.* **2021**, *174*, 655–662, doi:10.7326/M20-6976.
2. Niemi, M.E.K.; Karjalainen, J.; Liao, R.G.; Neale, B.M.; Daly, M.; Ganna, A.; Pathak, G.A.; Andrews, S.J.; Kanai, M.; Veerapen, K.; et al. Mapping the Human Genetic Architecture of COVID-19. *Nature* **2021**, *600*, 472–477, doi:10.1038/s41586-021-03767-x.
3. Muus, C.; Luecken, M.D.; Eraslan, G.; Sikkema, L.; Waghray, A.; Heimberg, G.; Kobayashi, Y.; Vaishnav, E.D.; Subramanian, A.; Smillie, C.; et al. Single-Cell Meta-Analysis of SARS-CoV-2 Entry Genes across Tissues and Demographics. *Nat. Med.* **2021**, *27*, 546–559, doi:10.1038/s41591-020-01227-z.
4. Ghamrawi, R.; Gunaratne, M. COVID-19 and Sex Differences. *Mayo Clin. Proc.* **2020**, *95*, 2189–2203, doi:10.1016/j.mayocp.2020.07.024.
5. Zhou, Y.; Chi, J.; Lv, W.; Wang, Y. Obesity and Diabetes as High-Risk Factors for Severe Coronavirus Disease 2019 (COVID-19). *Diabetes. Metab. Res. Rev.* **2021**, *37*, e3377, doi:10.1002/dmrr.3377.
6. Shokri, P.; Golmohammadi, S.; Noori, M.; Nejadghaderi, S.A.; Carson-Chahhoud, K.; Safiri, S. 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*, doi:https://doi.org/10.1002/rmv.2247.
7. Miftode, R.-S.; Costache, I.-I.; Cianga, P.; Petris, A.O.; Cianga, C.-M.; Maranduca, M.-A.; Miftode, I.-L.; Constantinescu, D.; Timpau, A.-S.; Crisan, A.; et al. The Influence of Socioeconomic Status on the Prognosis and Profile of Patients Admitted for Acute Heart Failure during COVID-19 Pandemic: Overestimated Aspects or a Multifaceted Hydra of Cardiovascular Risk Factors? *Healthc. (Basel, Switzerland)* **2021**, *9*, doi:10.3390/healthcare9121700.
8. Shelton, J.F.; Shastri, A.J.; Ye, C.; Weldon, C.H.; Filshtein-Sonmez, T.; Coker, D.; Symons, A.; Esparza-Gordillo, J.; Chubb, A.; Fitch, A.; et al. Trans-Ancestry Analysis Reveals Genetic and Nongenetic Associations with COVID-19 Susceptibility and Severity. *Nat. Genet.* **2021**, *53*, 801–808,

doi:10.1038/s41588-021-00854-7.

9. Choi, J.Y.; Smith, D.M. SARS-CoV-2 Variants of Concern. *Yonsei Med. J.* **2021**, *62*, 961–968, doi:10.3349/ymj.2021.62.11.961.
10. Ellinghaus, D.; Degenhardt, F.; Bujanda, L.; Buti, M.; Albillos, A.; Invernizzi, P.; Fernández, J.; Prati, D.; Baselli, G.; Asselta, R.; et al. Genomewide Association Study of Severe COVID-19 with Respiratory Failure. *N. Engl. J. Med.* **2020**, doi:10.1056/NEJMoa2020283.
11. Zeberg, H.; Pääbo, S. The Major Genetic Risk Factor for Severe COVID-19 Is Inherited from Neanderthals. *Nature* **2020**, *587*, 610–612, doi:10.1038/s41586-020-2818-3.
12. Pairo-Castineira, E.; Clohisey, S.; Klaric, L.; Bretherick, A.D.; Rawlik, K.; Pasko, D.; Walker, S.; Parkinson, N.; Fourman, M.H.; Russell, C.D.; et al. Genetic Mechanisms of Critical Illness in COVID-19. *Nature* **2021**, *591*, 92–98, doi:10.1038/s41586-020-03065-y.
13. Zhang, Q.; Bastard, P.; Liu, Z.; Le Pen, J.; Moncada-Velez, M.; Chen, J.; Ogishi, M.; Sabli, I.K.D.; Hodeib, S.; Korol, C.; et al. Inborn Errors of Type I IFN Immunity in Patients with Life-Threatening COVID-19. *Science (80-.).* **2020**, *370*, doi:10.1126/SCIENCE.ABD4570.
14. Cruz, R.; Lorenzo-salazar, J.M.; González-montelongo, R.; Gago-domínguez, M.; Porras, M.S.; Calderón, E.J.; Lattig, M.C.; Mancebo, E.; Yáñez, Z.; Zarate, R.; et al. Novel Genes and Sex Differences in COVID-19 Severity. **2022**, *31*, 3789–3806.
15. Wein, A.N.; McMaster, S.R.; Takamura, S.; Dunbar, P.R.; Cartwright, E.K.; Hayward, S.L.; McManus, D.T.; Shimaoka, T.; Ueha, S.; Tsukui, T.; et al. CXCR6 Regulates Localization of Tissue-Resident Memory CD8 T Cells to the Airways. *J. Exp. Med.* **2019**, *216*, 2748–2762, doi:10.1084/jem.20181308.
16. Duggal, P.; An, P.; Beaty, T.H.; Strathdee, S.A.; Farzadegan, H.; Markham, R.B.; Johnson, L.; O'Brien, S.J.; Vlahov, D.; Winkler, C.A. Genetic Influence of CXCR6 Chemokine Receptor Alleles on PCP-Mediated AIDS Progression among African Americans. *Genes Immun.* **2003**, *4*, 245–250, doi:10.1038/sj.gene.6363950.
17. Passam, A.M.; Sourvinos, G.; Krambovitis, E.; Miyakis, S.; Stavrianeas, N.; Zagoreos, I.; Spandidos, D.A. Polymorphisms of Cx(3)CR1 and CXCR6

Receptors in Relation to HAART Therapy of HIV Type 1 Patients. *AIDS Res. Hum. Retroviruses* **2007**, *23*, 1026–1032, doi:10.1089/aid.2006.0248.

18. Picton, A.C.P.; Paximadis, M.; Chaisson, R.E.; Martinson, N.A.; Tiemessen, C.T. CXCR6 Gene Characterization in Two Ethnically Distinct South African Populations and Association with Viraemic Disease Control in HIV-1-Infected Black South African Individuals. *Clin. Immunol.* **2017**, *180*, 69–79, doi:10.1016/j.clim.2017.04.006.
19. Limou, S.; Coulonges, C.; Herbeck, J.T.; van Manen, D.; An, P.; Le Clerc, S.; Delaneau, O.; Diop, G.; Taing, L.; Montes, M.; et al. Multiple-Cohort Genetic Association Study Reveals CXCR6 as a New Chemokine Receptor Involved in Long-Term Nonprogression to AIDS. *J. Infect. Dis.* **2010**, *202*, 908–915, doi:10.1086/655782.
20. Wei, Q.; Zhou, W.; Wang, W.; Gao, B.; Wang, L.; Cao, J.; Liu, Z.-P. Tumor-Suppressive Functions of Leucine Zipper Transcription Factor-like 1. *Cancer Res.* **2010**, *70*, 2942–2950, doi:10.1158/0008-5472.CAN-09-3826.
21. Promchan, K.; Natarajan, V. Leucine Zipper Transcription Factor-like 1 Binds Adaptor Protein Complex-1 and 2 and Participates in Trafficking of Transferrin Receptor 1. *PLoS One* **2020**, *15*, e0226298, doi:10.1371/journal.pone.0226298.
22. Downes, D.J.; Cross, A.R.; Hua, P.; Roberts, N.; Schwessinger, R.; Cutler, A.J.; Munis, A.M.; Brown, J.; Mielczarek, O.; de Andrea, C.E.; et al. Identification of LZTFL1 as a Candidate Effector Gene at a COVID-19 Risk Locus. *Nat. Genet.* **2021**, *53*, 1606–1615, doi:10.1038/s41588-021-00955-3.
23. Rescenko, R.; Peculis, R.; Briviba, M.; Ansone, L.; Terentjeva, A.; Litvina, H.D.; Birzniece, L.; Megnis, K.; Kolesova, O.; Rozentale, B.; et al. Replication of LZTFL1 Gene Region as a Susceptibility Locus for COVID-19 in Latvian Population. *Virol. Sin.* 2021, *36*, 1241–1244.
24. Angulo-Aguado, M.; Corredor-Orlandelli, D.; Carrillo-Martínez, J.C.; Gonzalez-Cornejo, M.; Pineda-Mateus, E.; Rojas, C.; Triana-Fonseca, P.; Contreras Bravo, N.C.; Morel, A.; Parra Abaunza, K.; et al. Association Between the LZTFL1 Rs11385942 Polymorphism and COVID-19 Severity in Colombian Population. *Front. Med.* **2022**, *9*, 910098, doi:10.3389/fmed.2022.910098.

25. Dai, Y.; Wang, J.; Jeong, H.H.; Chen, W.; Jia, P.; Zhao, Z. Association of CXCR6 with COVID-19 Severity: Delineating the Host Genetic Factors in Transcriptomic Regulation. *Hum. Genet.* **2021**, *140*, 1313–1328, doi:10.1007/s00439-021-02305-z.
26. Pereira, A.C.; Bes, T.M.; Velho, M.; Marques, E.; Jannes, C.E.; Valino, K.R.; Dinardo, C.L.; Costa, S.F.; Duarte, A.J.S.; Santos, A.R.; et al. Genetic Risk Factors and COVID-19 Severity in Brazil: Results from BRACOVID Study. *Hum. Mol. Genet.* **2022**, *31*, 3021–3031, doi:10.1093/hmg/ddac045.
27. Auton, A.; Abecasis, G.R.; Altshuler, D.M.; Durbin, R.M.; Bentley, D.R.; Chakravarti, A.; Clark, A.G.; Donnelly, P.; Eichler, E.E.; Flicek, P.; et al. A Global Reference for Human Genetic Variation. *Nature* **2015**, *526*, 68–74, doi:10.1038/nature15393.
28. Pena, S.D.J.; Santos, F.R.; Tarazona-Santos, E. Genetic Admixture in Brazil. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **2020**, *184*, 928–938, doi:10.1002/ajmg.c.31853.
29. Mistry, P.; Barmania, F.; Mellet, J.; Peta, K.; Strydom, A.; Viljoen, I.M.; James, W.; Gordon, S.; Pepper, M.S. SARS-CoV-2 Variants, Vaccines, and Host Immunity. *Front. Immunol.* **2021**, *12*, 809244, doi:10.3389/fimmu.2021.809244.

Variable	Level	Overall, n = 102	Mechanical ventilation			Death		
			No, n = 83	Yes, n = 19	p-value	No, n = 64	Yes, n = 38	p-value
Age (years)	-	60 (52, 68)	60 (52, 68)	65 (48, 68)	0.997	60 (52, 67)	66 (51, 70)	0.545
Hospitalization (days)	-	20 (9, 38)	18 (9, 35)	32 (10, 60)	0.125	19 (9, 41)	20 (10, 35)	0.890
	Female	46 (45%)	40 (48%)	6 (32%)		33 (52%)	13 (34%)	
Sex	Male	56 (55%)	43 (52%)	13 (68%)	0.189	31 (48%)	25 (66%)	0.089
	Not declared	5 (4.9%)	4 (4.8%)	1 (5.3%)		4 (6.2%)	1 (2.6%)	
	Yellow	3 (2.9%)	3 (3.6%)	0 (0%)		1 (1.6%)	2 (5.3%)	
	White	15 (14.7%)	10 (12.1%)	5 (26.3%)		5 (7.8%)	10 (26.3%)	
	Brown	75 (73.6%)	62 (74.7%)	13 (68.4%)		50 (78.1%)	25 (65.8%)	
Ethnicity	Black	4 (3.9%)	4 (4.8%)	0 (0%)	0.348	4 (6.3%)	0 (0%)	**0.015**
Systemic blood hypertension	Yes	57 (56%)	46 (55%)	11 (58%)	0.845	37 (58%)	20 (53%)	0.610
Neurological diseases	Yes	21 (21%)	17 (20%)	4 (21%)	>0.999	14 (22%)	7 (18%)	0.677
Hypothyroidism	Yes	12 (12%)	9 (11%)	3 (16%)	0.692	7 (11%)	5 (13%)	0.758
Cancer	Yes	3 (2.9%)	3 (3.6%)	0 (0%)	>0.999	1 (1.6%)	2 (5.3%)	0.554
Asthma	Yes	4 (3.9%)	2 (2.4%)	2 (11%)	0.157	2 (3.1%)	2 (5.3%)	0.627
HIV	Yes	7 (6.9%)	6 (7.2%)	1 (5.3%)	>0.999	5 (7.8%)	2 (5.3%)	>0.999
Alcoholism	Yes	5 (4.9%)	4 (4.8%)	1 (5.3%)	>0.999	3 (4.7%)	2 (5.3%)	>0.999
Smoker	Yes	5 (4.9%)	4 (4.8%)	1 (5.3%)	>0.999	4 (6.2%)	1 (2.6%)	0.648

Table 1: Clinical-epidemiological sample description. Median (interquartile range) or n (%) are presented.

Parameters		*CXCR6* rs2234355 (A/G)		*CXCR6* rs2234358 (G/T)		*LTZFL1* rs10490770 (C/T)	
		Mechanical ventilation	Death	Mechanical ventilation	Death	Mechanical ventilation	Death
	1/1	1	5	2	5	0	0
	1/2	5	5	14	24	3	6
	2/2	13	28	3	9	16	32
Case	HWE p-value	0.488	**0.001**	0.074	0.185	1.000	1.000
	1/1	6	2	13	10	0	0
	1/2	21	21	44	34	16	13
	2/2	56	41	26	20	67	51
Control	HWE p-value	0.078	1.000	0.504	0.609	1.000	1.000
Codominance p-value		0.951	**0.022**	0.235	0.604	0.567	0.720
Dominance p-value		0.936	0.311	0.156	0.409	-	-
Recessive p-value		0.753	0.057	0.555	0.732	-	-
Log-additive p-value		0.851	0.974	0.525	0.698	0.567	0.720

Table 2: Genetic association of *CXCR6* (rs2234355 and rs2234358) and *LZTFL1* (rs10490770) variants with the need for mechanical ventilation and death. HWE: Hardy-Weinberg equilibrium.

Model	Genotypes	*CXCR6* rs2234355 (A/G)			*CXCR6* rs2234458 (G/T)			*LZTFL1* rs10490770 (C/T)		
		n	mean (sd)	p-value	n	mean (sd)	p-value	n	mean (sd)	p-value
Codominance	1/1	69	30.67 (4.32)	0.504	29	26.10 (3.87)	0.738	0	0	0.771
	1/2	26	31.42 (5.83)		58	31.97 (5.02)		19	31.84 (6.8)	
	2/2	7	15.57 (3.22)		15	28.73 (8.67)		83	29.36 (3.75)	
Dominance	1/1	69	30.67 (4.32)	0.713	29	26.10 (3.87)	0.479	-	-	-
	1/2 + 2/2	33	28.06 (4.76)		73	31.30 (4.33)				
Recessive	1/1 + 1/2	95	30.87 (3.51)	0.242	87	30.01 (3.58)	0.892	-	-	-
	2/2	7	15.57 (3.22)		15	28.73 (8.67)				
Log-additive	0,1,2	-	-	0.446	-	-	0.673	-	-	0.771

Table 3: Association analysis between *CXCR6* rs2234355, *CXCR6* rs2234358, and *LZTFL1* rs10490770 variants and the hospitalization days in COVID-19 positive patients.

CAPÍTULO 5

Este capítulo aborda os resultados da análise de associação do polimorfismo rs12252 no gene *IFITM3* com a severidade da COVID-19, assim com a chance de óbito. O *IFITM3* foi o primeiro gene usado como candidato em estudos de associação com a doença. Este estudo gerou uma letter que apresentamos a seguir.

IFITM3 rs12252 polymorphism association with COVID-19 severity and mortality in a Brazilian sample: an update and a meta-analysis

João Locke Ferreira de Araújo[1]; Victória Frigério Bonifácio[2]; Lorena Medeiros Batista[2], Renata Eliane de Ávila[3], Renato Santana Aguiar [1,4], Luciana Bastos-Rodrigues[2,*]; Renan Pedra de Souza [1,*]

1 Grupo de Pesquisa em Bioestatística e Epidemiologia Molecular, Laboratório de Biologia Integrativa, Departamento de Genética, Ecologia e Evolução; Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

2 Departamento de Nutrição; Escola de Enfermagem, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

3 Hospital Eduardo de Menezes, Belo Horizonte, MG, Brazil

4 Instituto D'OR de Pesquisa e Ensino, Rio de Janeiro, RJ, Brazil

* Corresponding author: Renan P. Souza (renanpedra@gmail.com) or Luciana Bastos-Rodrigues (lu.bastosr@gmail.com). Universidade Federal de Minas Gerais. Av. Antônio Carlos, 6627 ICB – Pampulha, 31270901 – Belo Horizonte – Minas Gerais – Brazil. Phone: +553134092895.

Keywords: interferon; COVID-19 prognosis; SNP; polymorphisms; genetic association

Introduction

COVID-19 may present various symptoms. The prognosis is multifactorial and previous reports have suggested a role for the interferon-induced transmembrane protein 3 (*IFITM3*) gene. The encoded transmembrane protein restricts the cellular entry of pathogens, including influenza and ebola viruses [1,2]. SARS-CoV-2 infection susceptibility may also be regulated by *IFITM3* genetic variability [3]. *IFITM3* rs12252 polymorphism was the first to be explored in a candidate gene association study with a COVID-19 outcome. In a sample of 80 hospitalized Chinese patients, CC genotype carriers were 5.37 times more likely than CT or TT to develop severe COVID-19, defined by respiratory or other organ failures [4]. A recent meta-analysis found no association between the rs12252 polymorphism and ICU admission [5]. Here, we conducted an *IFITM3* rs12252 association study with severity and death in hospitalized Brazilian patients.

Methods

One hundred and two COVID-19 patients hospitalized at Hospital Eduardo de Menezes were enrolled between August and October 2020. COVID-19 molecular diagnosis was made using RT-qPCR. The UFMG Human Research Ethics Committee approved the project (number 31095820.4.0000.5149). *IFITM3* rs12252 genotypes were determined with a TaqMan probe (Thermo Fisher Scientific). Genotyping was performed blinded from clinical data and was repeated in 10% of the samples as a quality control procedure, with a 100% agreement rate.

Statistical analyses were performed in R version 4.0.2, considering a significance level of 0.05. Associations were assessed with Pearson's Chi-squared test or Fisher's exact test. Median age and days of hospitalizations were compared with Wilcoxon tests. We updated the search described in a previous systematic review [5] to evaluate the association of rs12252 with death. A meta-analysis combining odds ratios was carried out using the *metabin* function of the *meta* package.

Results

No significant differences or associations were found between outcomes and clinical variables (Table 1). The Hardy-Weinberg equilibrium was observed in all groups (case mechanical ventilation p-value: 0.53, control mechanical ventilation: 0.06, case death: 0.31, control death: 0.10). No *IFITM3* rs12252 association with severity or mortality was seen.

The updated search for the systematic review did not retrieve any additional studies exploring the same outcome. Therefore, we combined our findings with two previous studies [6,7]. No association was observed for the allelic (OR:1.37; 95%CI: 0.94-2.00) and the C-recessive model (OR:1.84; 95%CI: 0.27-12.53). However, the T recessive model was significant (OR:0.65; 95%CI: 0.43-0.99), assuming non-heterogeneity across the included studies (p=0.39) (Figure 1).

Discussion

Association results of the rs12252 polymorphism with COVID-19 outcomes have been heterogeneous. While the first report was significant [4], Gomez and collaborators [8], followed by two other studies [7,9], did not find increased odds of requiring intensive care unit treatment in CC carriers. While our original finding corroborates the lack of association, we observed reduced death odds in TT carriers compared to CC and CT subjects using meta-analysis. It is crucial to point out that the T allele is rare, and its homozygosity was not found in our sample. Replications in larger and ethnically diverse samples are warranted to understand further the *IFITM3* rs12252 role in COVID-19 prognosis

Conflict of interest: The authors did not show any conflict of interest.

Acknowledgment: None.

Funding

We acknowledge support from the Rede Corona-ômica BR MCTI/FINEP affiliated with RedeVírus/MCTI (01.20.0029.000462/20 404096/2020-4; 1227/21 01.22.0074.00); Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (315592/2021-4); Financiadora de Estudos e Projetos - FINEP (0494/20 01.20.0026.00; 1228/21 01.22.0082.00; 1139/20 01.20.0076.00); Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES (Finance Code 001) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais—FAPEMIG (APQ-00475-20)

Reference

1. Prabhu, S.S.; Chakraborty, T.T.; Kumar, N.; Banerjee, I. Association between IFITM3 Rs12252 Polymorphism and Influenza Susceptibility and Severity: A Meta-Analysis. *Gene* **2018**, *674*, 70–79, doi:10.1016/j.gene.2018.06.070.

2. Harmand, T.J.; Pattabiraman, V.R.; Bode, J.W. Chemical Synthesis of the Highly Hydrophobic Antiviral Membrane-Associated Protein IFITM3 and Modified Variants. *Angew. Chem. Weinheim. Bergstr. Ger.* **2017**, *129*, 12813–12817, doi:10.1002/ange.201707554.

3. Xu, F.; Wang, G.; Zhao, F.; Huang, Y.; Fan, Z.; Mei, S.; Xie, Y.; Wei, L.; Hu, Y.; Wang, C.; et al. IFITM3 Inhibits SARS-CoV-2 Infection and Is Associated with COVID-19 Susceptibility. *Viruses* **2022**, *14*, doi:10.3390/v14112553.

4. Zhang, Y.; Zhang, Y.; Qin, L.; Zhao, Y.; Zhang, P.; Zhang, P.; Xu, B.; Li, K.; Liang, L.; Zhang, C.; et al. Interferon-Induced Transmembrane Protein 3 Genetic Variant Rs12252-C Associated with Disease Severity in Coronavirus Disease 2019. *J. Infect. Dis.* **2020**, *222*, 34–37, doi:10.1093/infdis/jiaa224.

5. de Araújo, J.L.F.; Menezes, D.; de Aguiar, R.S.; de Souza, R.P. IFITM3, FURIN, ACE1, and TNF-α Genetic Association With COVID-19 Outcomes: Systematic Review and Meta-Analysis. *Front. Genet.* 2022, *13*, 775246.

6. Alghamdi, J.; Alaamery, M.; Barhoumi, T.; Rashid, M.; Alajmi, H.; Aljasser, N.; Alhendi, Y.; Alkhalaf, H.; Alqahtani, H.; Algablan, O.; et al. Interferon-Induced Transmembrane Protein-3 Genetic Variant Rs12252 Is Associated with COVID-19 Mortality. *Genomics* **2021**, *113*, 1733–1741, doi:10.1016/j.ygeno.2021.04.002.

7. Cuesta-Llavona, E.; Albaiceta, G.M.; García-Clemente, M.; Duarte-Herrera, I.D.; Amado-Rodríguez, L.; Hermida-Valverde, T.; Enríquez-Rodriguez, A.I.; Hernández-González, C.; Melón, S.; Alvarez-Argüelles, M.E.; et al. Association between the Interferon-Induced Transmembrane Protein 3 Gene (IFITM3) Rs34481144 / Rs12252 Haplotypes and COVID-19. *Curr. Res. Virol. Sci.* **2021**, *2*, 100016, doi:10.1016/j.crviro.2021.100016.

8. Gómez, J.; Albaiceta, G.M.; Cuesta-Llavona, E.; García-Clemente, M.; López-Larrea, C.; Amado-Rodríguez, L.; López-Alonso, I.; Melón, S.; Alvarez-Argüelles, M.E.; Gil-Peña, H.; et al. The Interferon-Induced Transmembrane Protein 3 Gene (IFITM3) Rs12252 C Variant Is Associated with COVID-19. *Cytokine* **2021**, *137*, 155354, doi:10.1016/j.cyto.2020.155354.

9. Schönfelder, K.; Breuckmann, K.; Elsner, C.; Dittmer, U.; Fistera, D.; Herbstreit, F.; Risse, J.; Schmidt, K.; Sutharsan, S.; Taube, C.; et al. The Influence of IFITM3 Polymorphisms on Susceptibility to SARS-CoV-2 Infection and Severity of COVID-19. *Cytokine* **2021**, *142*, 155492, doi:10.1016/j.cyto.2021.155492.

Table 1: Evaluation of factors associated with severity and mortality outcomes. No significant effects were observed.

Phenotype	Mechanical ventilation		p-value	Death		p-value
	No, n = 83	Yes, n = 19		No, n = 64	Yes, n = 38	
Age (years)	60 (52,68)	65 (48,68)	0.997	60 (52, 67)	66 (51,70)	0.545
Hospitalization (days)	18 (9,35)	32 (10,60)	0.125	19 (9,41)	20 (10,35)	0.890
Sex female	40 (48%)	6 (32%)		33 (52%)	13 (34%)	
Sex male	43 (52%)	13 (68%)	0.189	31 (48%)	25 (66%)	0.089
Systemic arterial hypertension	46 (55%)	11 (58%)	0.845	37 (58%)	20 (53%)	0.610
Neurological diseases	17 (20%)	4 (21%)	>0.999	14 (22%)	7 (18%)	0.677
Hypothyroidism	9 (11%)	3 (16%)	0.692	7 (11%)	5 (13%)	0.758
Cancer	3 (3.6%)	0 (0%)	>0.999	1 (1.6%)	2 (5.3%)	0.554
Asthma	2 (2.4%)	2 (11%)	0.157	2 (3.1%)	2 (5.3%)	0.627
HIV infection	6 (7.2%)	1 (5.3%)	>0.999	5 (7.8%)	2 (5.3%)	>0.999
Alcoholism	4 (4.8%)	1 (5.3%)	>0.999	3 (4.7%)	2 (5.3%)	>0.999
Smoker	4 (4.8%)	1 (5.3%)	>0.999	4 (6.2%)	1 (2.6%)	0.648
rs12252 T/T	53 (64%)	11 (58%)		40 (62%)	24 (63%)	
rs12252 T/C	30 (36%)	8 (42%)	0.628	24 (38%)	14 (37%)	0.947

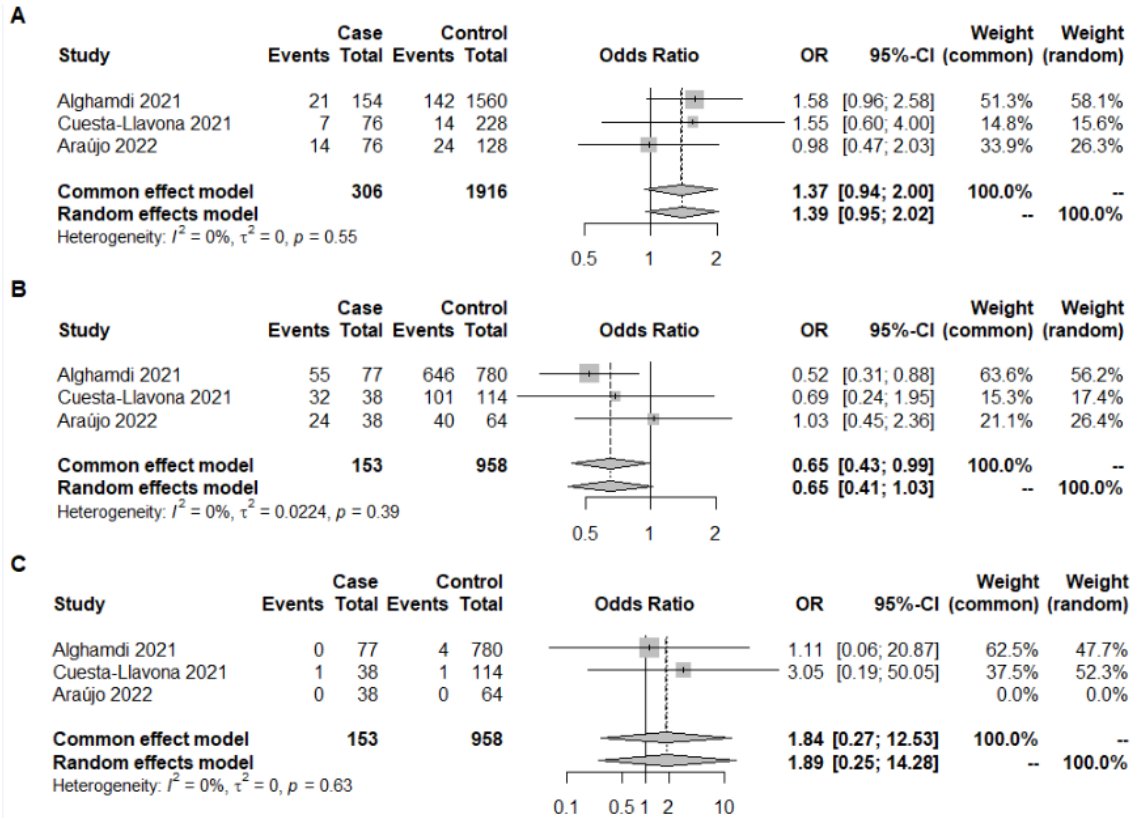


Figure 1- *IFITM3* rs12252 association with mortality under the (A) C-allele model. (B) T-recessive model. (C) C-recessive model. The T-recessive model was significant.

DISCUSSÃO GERAL

O estudo da genética humana e a interação vírus hospedeiro sempre foram foco para a descoberta de abordagens tanto terapêuticas quanto para a intervenção no controle da disseminação de doenças infecciosas. Objetivamos nesta tese investigar possíveis candidatos a marcadores moleculares para prever risco de gravidade para a COVID-19 e a susceptibilidade à infecção pelo SARS-CoV-2. Através de uma revisão sistemática foi possível conhecermos como os estudos de associação envolvendo estes dois desfechos foram conduzidos no primeiro ano de pandemia, e dessa forma focarmos nosso olhar para marcadores promissores.

Desde então, muitos outros trabalhos envolvendo genes candidatos foram publicados, muitos deles devido a estudos em escala genômica e tentativas de replicações destes resultados ou hipóteses com boa ou razoável plausibilidade biológica. Seguindo o mesmo caminho fomos capazes de combinar efeitos através de metanálise de alguns destes marcadores, tendo em vista que os estudos com os mesmos polimorfismos seguiram sendo replicados, possibilitando esta abordagem.

Ao logo desta tese, conseguimos aplicar a metodologia de metanálise em outros estudos originais, além do capítulo 2 (uma revisão sistemática com metanálise), onde além de combinarmos efeitos entre coortes (capítulo 3), foi possível combinarmos efeitos dos nossos estudos originais com os encontrados na literatura (capítulos 3 e 5). As metanálises são consideradas padrão ouro de evidência cientifica, e a exploração de seu uso na pesquisa serve de grande apoio para a qualidade da interpretação de resultados encontrados.

Ressaltamos aqui a importância do cuidado com a qualidade dos estudos inclusos em metanálises, além da importância de se publicar resultados negativos de bons estudos conduzidos. O instrumento Q-Genie foi fundamental para aumentar a robustez de nosso trabalho, permitindo incluir estudos de associação com avaliação acima de moderada e excluindo estudos pobres na combinação de efeitos. Mas este instrumento serve para a avaliação de estudos de associação genética, sendo fundamentado no STREGA. Existem outros instrumentos de avaliação para outros tipos de desenhos experimentais que gostaríamos de encorajar ao uso, antes de uma revisão sistemática, seguida de metanálise. O Joanna Briggs Institute (https://jbi.global/) por exemplo, é uma organização internacional de pesquisa que oferece instrumentos de avaliação cientométrica para diversos tipos de estudos. A

tomada de decisão baseada em evidências deve ser vista como prioridade na academia, e o trabalho para que o resultado observado seja o mais próximo da realidade possível depende de um desenho experimental bem estruturado e conduzido.

Uma vez compreendendo o atual cenário de biomarcadores, consideramos interessante avaliar os genes *ACE2* e *TMPRSS2* como primeiro passo, já que ambos representam loci candidatos clássicos considerando os mecanismos de entrada na célula pelo vírus, além do polimorfismo rs4646994 no gene *ACE1,* tendo em vista as evidências crescentes de que este marcador possa contribuir para a gravidade na COVID-19. Colaboramos com a publicação de um trabalho dosando a expressão destes genes com o laboratório de virologia molecular, na UFRJ, que é citado no anexo 1.

Consideramos no início polimorfismos nos genes *ACE2* e *TMPRSS2*, mas a baixa frequência destas variantes e nossa coorte de tamanho limitado na vez procurar por outros biomarcadores de equivalente plausibilidade biológica.

Algumas questões precisam ser consideradas. Primeiramente, o volume de estudos de associação apresentados na primeira revisão ainda era pequeno para que fosse possível tomar decisões em cima de algum resultado. Em 2020, apenas 20 estudos foram publicados. Na mesma linha, apesar das metanálises representarem um padrão ouro de evidência, o número reduzido de trabalhos além da heterogeneidade no método dos estudos utilizados na análise prejudica a confiabilidade do resultado. Mais estudos precisariam ser replicados com esses marcadores e de uma forma mais homogênea para que os efeitos possam ser combinados da melhor forma possível.

Em 2022, foram publicadas outras revisões sistemáticas e metanálises com estudos de associação. Uma revisão publicada na revista Gene em junho deste ano encontrou 60 estudos de associação, entre estudos com genes candidatos ou estudos em escala genômica.[135]. Outras revisões abordando genética de associação também foram publicadas em 2021[136–138]. As metanálises conduzidas para a deleção no *ACE1* com o prognóstico e a suscetibilidade da COVID-19 no estudo apresentado no capítulo 2 desta tese são as mais completas até o momento.

Nosso tamanho amostral no estudo de expressão gênica foi limitado. O ensaio de expressão de alguns indivíduos, assim como a genotipagem, não puderam ser concluídos devida principalmente a qualidade da amostra coletada, muitas vezes prejudicada devido ao tempo de armazenagem ou ao amplo uso dela para outros

experimentos. Durante esta tese, muitas hipóteses foram levantadas, assim como outros biomarcadores para serem trabalhados tanto em estudos de associação envolvendo polimorfismos genéticos, quanto a ensaios de expressão. Os genes *FURIM*, *CTSL* e *DPP4*, foram cogitados para ensaios de expressão para análise da associação dos níveis de expressão no prognóstico clínico da COVID-19. Contudo enfrentamos dificuldades na padronização dos ensaios, e optamos para deixar estes experimentos para o futuro.

Tambem a exploração de polimorfismos no complexo *HLA* seria interessante em amostras nacionais, dada a enorme miscigenação da população brasileira e a pluralidade de alelos no *HLA*. Seria uma possibilidade de continuidade de trabalho, considerando a importância dos estudos de replicação e a possiblidade de se encontrar alelos múltiplos com frequências consideráveis na população brasileira. Uma revisão sistemática publicada em 2022 identificou 36 estudos abordando polimorfismos no HLA na susceptibilidade, severidade e chance de óbito na COVID-19[139]. Alem disso, um estudo com amostras brasileiras foi publicado[140]. Contudo este procedimento levaria um tempo maior devida à necessidade de protocolos de sequenciamento, por se tratar de loci multialélicos. É extremamente importante explorar outros marcadores que possam contribuir para a explicação da enorme heterogeneidade do desfecho clínico da COVID-19, além de novas abordagens de análise. Nossa amostragem é composta de indivíduos infectados com vírus de linhagens de 2020, e nos últimos 2 anos as variantes de preocupação tem dominado os casos de infectados, e muitos estudos já os relacionam com gravidade da COVID-19, assim como maiores chances de transmissão e até mesmo resistência a vacinas. A replicação destes estudos em pacientes infectados com outras cepas e compará-los com os resultados já observados em pacientes infectados em 2020 é um passo considerável no estudo da heterogeneidade da doença. Tambem é importante considerar o processo de imunização que começou a se fortalecer no segundo semestre de 2021 e que com certeza influenciou no controle da pandemia. Por fim, um olhar atento a deleção de 287pb no gene *ACE1* mostra-se fundamental, considerando os resultados observados em estudos de caso-controle até o momento, indicando este polimorfismo como forte biomarcador de gravidade na COVID-19.

REFERÊNCIAS

1. Naqvi, A. A. T. *et al.* Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1866**, 165878 (2020).

2. Malik, Y. A. Properties of Coronavirus and SARS-CoV-2. *Malays. J. Pathol.* **42**, 3–11 (2020).

3. V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* **19**, 155–170 (2021).

4. Tang, X. *et al.* On the origin and continuing evolution of SARS-CoV-2. 1012–1023 (2020). doi:10.1093/nsr/nwaa036

5. Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet (London, England)* **395**, 565–574 (2020).

6. Tobin, M. J. Basing Respiratory Management of COVID-19 on Physiological Principles. *American journal of respiratory and critical care medicine* **201**, 1319–1320 (2020).

7. Mönkemüller, K., Fry, L. & Rickes, S. COVID-19, coronavirus, SARS-CoV-2 and the small bowel. *Rev. Esp. enfermedades Dig. organo Of. la Soc. Esp. Patol. Dig.* **112**, 383–388 (2020).

8. Rabaan, A. A. *et al.* SARS-CoV-2, SARS-CoV, and MERS-COV: A comparative overview. *Le Infez. Med.* **28**, 174–184 (2020).

9. Chen, N. *et al.* Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet (London, England)* **395**, 507–513 (2020).

10. Xu, X. & Chen, P. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. **63**, 457–460 (2020).

11. Madabhavi, I., Sarkar, M. & Kadakol, N. COVID-19: a review. *Monaldi Arch. chest Dis. = Arch. Monaldi per le Mal. del torace* **90**, (2020).

12. Izurieta, H. S. *et al.* Natural History of Coronavirus Disease 2019: Risk Factors for Hospitalizations and Deaths Among >26 Million US Medicare Beneficiaries. *J. Infect. Dis.* (2020). doi:10.1093/infdis/jiaa767

13. 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* **15**, e0241265 (2020).

14. Gidari, A. *et al.* Cross-neutralization of SARS-CoV-2 B.1.1.7 and P.1 variants in vaccinated, convalescent and P.1 infected. *J. Infect.* **83**, 467–472 (2021).

15. Conti, P. *et al.* The British variant of the new coronavirus-19 (Sars-Cov-2) should not create a vaccine problem. *Journal of biological regulators and homeostatic agents* **35**, (2021).

16. Li, Q. *et al.* The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell* **182**, 1284-1294.e9 (2020).

17. Naqvi, A. A. T. *et al.* Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. *Biochim. Biophys. acta. Mol. basis Dis.* **1866**, 165878 (2020).

18. Bittar, C. *et al.* The Emergence of the New P.4 Lineage of SARS-CoV-2 With Spike L452R Mutation in Brazil. *Front. public Heal.* **9**, 745310 (2021).

19. Konings, F. *et al.* SARS-CoV-2 Variants of Interest and Concern naming scheme conducive for global discourse. *Nat. Microbiol.* **6**, 821–823 (2021).

20. Nyberg, T. *et al.* Risk of hospital admission for patients with SARS-CoV-2 variant B.1.1.7: cohort analysis. *BMJ* **373**, n1412 (2021).

21. Ramundo, M. S., Jesus, J. G. De, Andrade, P. S. & Coletti, T. M. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. **821**, 815–821 (2021).

22. Claro, I. M. *et al.* Local Transmission of SARS-CoV-2 Lineage B.1.1.7, Brazil, December 2020. *Emerg. Infect. Dis.* **27**, 970–972 (2021).

23. Tillett, R. L. *et al.* Genomic evidence for reinfection with SARS-CoV-2: a case study. *Lancet. Infect. Dis.* **21**, 52–58 (2021).

24. Nonaka, C. K. V *et al.* Genomic Evidence of SARS-CoV-2 Reinfection Involving E484K Spike Mutation, Brazil. *Emerg. Infect. Dis.* **27**, 1522–1524 (2021).

25. Tao, K. *et al.* The biological and clinical significance of emerging SARS-CoV-2 variants. *Nat. Rev. Genet.* **22**, 757–773 (2021).

26. Pinto, D. *et al.* SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. doi:10.1038/s41586-021-03944-y

27. Menezes, D., Fonseca, P. L. C., de Araújo, J. L. F. & Souza, R. P. de. SARS-CoV-2 Genomic Surveillance in Brazil: A Systematic Review with Scientometric Analysis. *Viruses* **14**, (2022).

28. Lamarca, A. P. *et al.* Genomic Surveillance Tracks the First Community Outbreak of the SARS-CoV-2 Delta (B.1.617.2) Variant in Brazil. *Journal of virology* **96**, e0122821 (2022).

29. Dos Santos, J. et al. (2021), First Reported Cases of SARS-CoV-2 Sublineage B.1.617.2 In Brazil: An Outbreak in a Ship and Alert for Spread<

https://virological.org/t/first-reported-cases-of-sars-cov-2-sub-lineage-b-1-617-2-in-brazil-an-outbreak-in-a-ship-and-alert-for-spread/706> accessed 30 Nov 2021

30. Giovanetti, M. *et al.* Replacement of the Gamma by the Delta variant in Brazil: Impact of lineage displacement on the ongoing pandemic. *Virus Evol.* **8**, veac024 (2022).

31. Alcantara, L. C. J. *et al.* SARS-CoV-2 epidemic in Brazil: how the displacement of variants has driven distinct epidemic waves. *Virus Res.* **315**, 198785 (2022).

32. Viana, R. *et al.* Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature* (2022). doi:10.1038/s41586-022-04411-y

33. Takashita, E. *et al.* Efficacy of Antiviral Agents against the SARS-CoV-2 Omicron Subvariant BA.2. *The New England journal of medicine* **386**, 1475–1477 (2022).

34. Cao, Y. *et al.* BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. *Nature* **608**, 593–602 (2022).

35. Scarpa, F. *et al.* Genetic and Structural Data on the SARS-CoV-2 Omicron BQ.1 Variant Reveal Its Low Potential for Epidemiological Expansion. *Int. J. Mol. Sci.* **23**, (2022).

36. Smit, M., Marinosci, A., Agoritsas, T. & Calmy, A. Prophylaxis for COVID-19: a systematic review. *Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis.* **27**, 532–537 (2021).

37. Gottlieb, R. L. *et al.* Effect of Bamlanivimab as Monotherapy or in Combination With Etesevimab on Viral Load in Patients With Mild to Moderate COVID-19: A Randomized Clinical Trial. *JAMA* **325**, 632–644 (2021).

38. Cohen, M. S. *et al.* Effect of Bamlanivimab vs Placebo on Incidence of COVID-19 Among Residents and Staff of Skilled Nursing and Assisted Living Facilities A Randomized Clinical Trial. **46225**, 46–55 (2021).

39. Oran, D. P. & Topol, E. J. The Proportion of SARS-CoV-2 Infections That Are Asymptomatic: A Systematic Review. *Ann. Intern. Med.* **174**, 655–662 (2021).

40. Wiersinga, W. J., Rhodes, A., Cheng, A. C., Peacock, S. J. & Prescott, H. C. Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. *JAMA* **324**, 782–793 (2020).

41. Emery, J. C. *et al.* The contribution of asymptomatic SARS-CoV-2 infections to transmission on the Diamond Princess cruise ship. *Elife* **9**, (2020).

42. Gao, Z. *et al.* A systematic review of asymptomatic infections with COVID-19. *J. Microbiol. Immunol. Infect.* **54**, 12–16 (2021).

43. Krieger, E. *et al.* Immunological determinants of clinical outcomes in COVID-19: A quantitative perspective. 1–36 (2020).

44. Lei, Q. *et al.* Antibody dynamics to SARS-CoV-2 in asymptomatic COVID-19 infections. *Allergy* **76**, 551–561 (2021).

45. He, J., Tao, H., Yan, Y., Huang, S. Y. & Xiao, Y. Molecular mechanism of evolution and human infection with SARS-CoV-2. *Viruses* **12**, (2020).

46. Yesudhas, D., Srivastava, A. & Gromiha, M. M. COVID-19 outbreak: history, mechanism, transmission, structural studies and therapeutics. *Infection* **49**, 199–213 (2021).

47. Kowalik, M. M. *et al.* COVID-19 - toward a comprehensive understanding of the disease. *Cardiol. J.* **27**, 99–114 (2020).

48. Lee, I.-C., Huo, T.-I. & Huang, Y.-H. Gastrointestinal and liver manifestations in patients with COVID-19. *J. Chin. Med. Assoc.* **83**, 521–523 (2020).

49. Carrillo-Larco, R. M. & Altez-Fernandez, C. Anosmia and dysgeusia in COVID-19: A systematic review. *Wellcome open research* **5**, 94 (2020).

50. Izcovich, A. *et al.* Prognostic factors for severity and mortality in patients infected with COVID-19: A systematic review. *PLoS One* **15**, e0241955 (2020).

51. Berlin, D. A., Gulick, R. M. & Martinez, F. J. Severe COVID-19. *N. Engl. J. Med.* **383**, 2451–2460 (2020).

52. Ding, S. & Liang, T. J. Is SARS-CoV-2 Also an Enteric Pathogen With Potential Fecal-Oral Transmission? A COVID-19 Virological and Clinical Review. *Gastroenterology* **159**, 53–61 (2020).

53. Gu, J., Han, B. & Wang, J. COVID-19: Gastrointestinal Manifestations and Potential Fecal-Oral Transmission. *Gastroenterology* **158**, 1518–1519 (2020).

54. Pollitt, K. J. G. *et al.* COVID-19 vulnerability: the potential impact of genetic susceptibility and airborne transmission. 1–7 (2020).

55. Setti, L. *et al.* Airborne Transmission Route of COVID-19: Why 2 Meters/6 Feet of Inter-Personal Distance Could Not Be Enough. *International journal of environmental research and public health* **17**, (2020).

56. Morawska, L. *et al.* How can airborne transmission of COVID-19 indoors be minimised? *Environment international* **142**, 105832 (2020).

57. Güner, R., Hasanoğlu, I. & Aktaş, F. COVID-19: Prevention and control measures in community. *Turkish J. Med. Sci.* **50**, 571–577 (2020).

58. Wang, J., Pan, L., Tang, S., Ji, J. S. & Shi, X. Mask use during COVID-19: A risk adjusted strategy. *Environ. Pollut.* **266**, 115099 (2020).

59. Hussain, M. *et al.* Structural variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein. *J. Med. Virol.* (2020). doi:10.1002/jmv.25832

60. Bourgonje, A. R. *et al.* Angiotensin-converting enzyme 2 (ACE2), SARS-CoV-2 and the pathophysiology of coronavirus disease 2019 (COVID-19). *J. Pathol.* **251**, 228–248 (2020).

61. Phan, T. Genetic diversity and evolution of SARS-CoV-2. *Infect. Genet. Evol.* **81**, (2020).

62. Verdecchia, P., Cavallini, C., Spanevello, A. & Angeli, F. The pivotal link between ACE2 deficiency and SARS-CoV-2 infection. *Eur. J. Intern. Med.* **76**, 14–20 (2020).

63. Muus, C. *et al. Single-cell meta-analysis of SARS-CoV-2 entry genes across tissues and demographics*. *Nature Medicine* **27**, (2021).

64. Al-mulla, F. *et al.* A comprehensive germline variant and expression analyses of ACE2, TMPRSS2 and SARS-CoV-2 activator FURIN genes from the Middle East : Combating SARS-CoV-2 with precision medicine. (2020).

65. Rossi, Á. D. *et al.* Association between ACE2 and TMPRSS2 nasopharyngeal expression and COVID-19 respiratory distress. *Sci. Rep.* **11**, 9658 (2021).

66. Hu, X. *et al.* Discovery of TMPRSS2 inhibitors from virtual screening. *bioRxiv : the preprint server for biology* (2021). doi:10.1101/2020.12.28.424413

67. Wang, Z. *et al.* Significance of the TMPRSS2:ERG gene fusion in prostate cancer. *Mol. Med. Rep.* **16**, 5450–5458 (2017).

68. Brojakowska, A., Narula, J., Shimony, R. & Bander, J. Clinical Implications of SARS-Cov2 Interaction with Renin Angiotensin System. *J. Am. Coll. Cardiol.* (2020). doi:10.1016/j.jacc.2020.04.028

69. Noori, M. *et al.* How SARS-CoV-2 might affect potassium balance via impairing epithelial sodium channels? *Mol. Biol. Rep.* **48**, 6655–6661 (2021).

70. Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280.e8 (2020).

71. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292.e6 (2020).

72. Gomes, C. P. *et al.* Cathepsin L in COVID-19: From Pharmacological Evidences to Genetics. *Front. Cell. Infect. Microbiol.* **10**, 1–10 (2020).

73. Wilk, A. J. *et al.* A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat. Med.* **26**, 1070–1076 (2020).

74. Vargas-Alarcón, G., Posadas-Sánchez, R. & Ramírez-Bello, J. Variability in genes related to SARS-CoV-2 entry into host cells (ACE2, TMPRSS2, TMPRSS11A, ELANE, and CTSL) and its potential use in association studies. *Life Sci.* **260**, 118313 (2020).

75. Sachan, V. *et al.* HIV-induced neuroinflammation: impact of PAR1 and PAR2 processing by Furin. *Cell Death Differ.* **26**, 1942–1954 (2019).

76. Kouretova, J. *et al.* Effects of NS2B-NS3 protease and furin inhibition on West Nile and Dengue virus replication. *J. Enzyme Inhib. Med. Chem.* **32**, 712–721 (2017).

77. Vankadari, N. Structure of Furin Protease Binding to SARS-CoV-2 Spike Glycoprotein and Implications for Potential Targets and Virulence. *J. Phys. Chem. Lett.* **11**, 6655–6663 (2020).

78. Hoffmann, M., Kleine-Weber, H. & Pöhlmann, S. A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Mol. Cell* **78**, 779-784.e5 (2020).

79. Ryabkova, V. A., Churilov, L. P. & Shoenfeld, Y. Influenza infection, SARS, MERS and COVID-19: Cytokine storm - The common denominator and the lessons to be learned. *Clin. Immunol.* **223**, 108652 (2021).

80. Hu, B., Huang, S. & Yin, L. The cytokine storm and COVID-19. *J. Med. Virol.* **93**, 250–256 (2021).

81. Luo, X.-H., Zhu, Y., Mao, J. & Du, R.-C. T cell immunobiology and cytokine storm of COVID-19. *Scand. J. Immunol.* **93**, e12989 (2021).

82. Xiang, Z. *et al.* Glucocorticoids improve severe or critical COVID-19 by activating ACE2 and reducing IL-6 levels. *Int. J. Biol. Sci.* **16**, 2382–2391 (2020).

83. Anka, A. U. *et al.* Coronavirus disease 2019 (COVID-19): An overview of the immunopathology, serological diagnosis and management. *Scand. J. Immunol.* **93**, e12998 (2021).

84. Updated, L. BEST (Biomarkers, EndpointS , and other Tools) Resource. (2021).

85. Smieszek, S. P., Przychodzen, B. P., Polymeropoulos, V. M., Polymeropoulos, C. M. & Polymeropoulos, M. H. Assessing the potential correlation of polymorphisms in the IL6R with relative IL6 elevation in severely ill COVID-19 patients'. *Cytokine* 155662 (2021). doi:10.1016/j.cyto.2021.155662

86. Karcioglu Batur, L. & Hekim, N. Correlation between interleukin gene polymorphisms and current prevalence and mortality rates due to novel coronavirus disease 2019 (COVID-2019) in 23 countries. *J. Med. Virol.* **93**, 5853–5863 (2021).

87. Akbari, H. *et al.* The role of cytokine profile and lymphocyte subsets in the severity of coronavirus disease 2019 (COVID-19): A systematic review and meta-analysis. *Life Sci.* **258**, 118167 (2020).

88. Rose-John, S. IL-6 trans-signaling via the soluble IL-6 receptor: importance for the pro-inflammatory activities of IL-6. *Int. J. Biol. Sci.* **8**, 1237–1247 (2012).

89. Rose-John, S. The Soluble Interleukin 6 Receptor: Advanced Therapeutic Options in Inflammation. *Clin. Pharmacol. Ther.* **102**, 591–598 (2017).

90. Tanaka, T., Narazaki, M. & Kishimoto, T. IL-6 in inflammation, immunity, and disease. *Cold Spring Harb. Perspect. Biol.* **6**, a016295 (2014).

91. Qu, R. *et al.* Platelet-to-lymphocyte ratio is associated with prognosis in patients with coronavirus disease-19. *J. Med. Virol.* **92**, 1533–1541 (2020).

92. Coomes, E. A. & Haghbayan, H. Interleukin-6 in COVID-19: A systematic review and meta-analysis. *Rev. Med. Virol.* **30**, 1–9 (2020).

93. Rostami, M. & Mansouritorghabeh, H. D-dimer level in COVID-19 infection: a systematic review. *Expert Rev. Hematol.* **13**, 1265–1275 (2020).

94. Gómez-Mesa, J. E., Galindo-Coral, S., Montes, M. C. & Muñoz Martin, A. J. Thrombosis and Coagulopathy in COVID-19. *Curr. Probl. Cardiol.* **46**, 100742 (2021).

95. Sastry, S., Cuomo, F. & Muthusamy, J. COVID-19 and thrombosis: The role of hemodynamics. *Thromb. Res.* **212**, 51–57 (2022).

96. Düz, M. E., Balcı, A. & Menekşe, E. D-dimer levels and COVID-19 severity: Systematic Review and Meta-Analysis. *Tuberk. Toraks* **68**, 353–360 (2020).

97. Turan, O. *et al.* Clinical characteristics, prognostic factors, and maternal and neonatal outcomes of SARS-CoV-2 infection among hospitalized pregnant women: A systematic review. *Int. J. Gynaecol. Obstet. Off. organ Int. Fed. Gynaecol. Obstet.* **151**, 7–16 (2020).

98. Katneni, U. K. *et al.* Coagulopathy and Thrombosis as a Result of Severe COVID-19 Infection: A Microvascular Focus. *Thromb. Haemost.* **120**, 1668–1679 (2020).

99. Chan, N. C. & Weitz, J. I. COVID-19 coagulopathy, thrombosis, and bleeding. *Blood* **136**, 381–383 (2020).

100. Connors, J. M. & Levy, J. H. COVID-19 and its implications for thrombosis and anticoagulation. *Blood* **135**, 2033–2040 (2020).

101. Billett, H. H. *et al.* Anticoagulation in COVID-19: Effect of Enoxaparin, Heparin, and Apixaban on Mortality. *Thromb. Haemost.* **120**, 1691–1699 (2020).

102. Suryamohan, K. *et al.* Human ACE2 receptor polymorphisms and altered susceptibility to SARS-CoV-2. *Commun. Biol.* **4**, (2021).

103. Delanghe, J. R., Speeckaert, M. M. & Buyzere, M. L. De. The host’s angiotensin-converting enzyme polymorphism may explain epidemiological findings in COVID-19 infections. *Clin. Chim. Acta* (2020). doi:10.1016/j.cca.2020.03.031

104. Zipeto, D., Palmeira, J. da F., Argañaraz, G. A. & Argañaraz, E. R. ACE2/ADAM17/TMPRSS2 Interplay May Be the Main Risk Factor for COVID-19. *Front. Immunol.* **11**, 1–10 (2020).

105. Onabajo, O. O. *et al.* Interferons and viruses induce a novel truncated ACE2 isoform and not the full-length SARS-CoV-2 receptor. *Nat. Genet.* **52**, 1283–1293 (2020).

106. Littera, R. *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.* **11**, 605688 (2020).

107. Di Maria, E., Latini, A., Borgiani, P. & Novelli, G. Genetic variants of the human host influencing the coronavirus-associated phenotypes (SARS, MERS and COVID-19): rapid systematic review and field synopsis. *Hum. Genomics* **14**, 30 (2020).

108. Amodio, E. *et al.* SARS-CoV-2 Viral Load, IFNλ Polymorphisms and the Course of COVID-19: An Observational Study. *J. Clin. Med.* **9**, 3315 (2020).

109. Ragia, G. & Manolopoulos, V. G. Assessing COVID-19 susceptibility through analysis of the genetic and epigenetic diversity of ACE2-mediated SARS-CoV-2 entry. *Pharmacogenomics* **21**, 1311–1329 (2020).

110. Bestle, D. *et al.* TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci. alliance* **3**, (2020).

111. Yang, X. *et al.* FURIN Expression in Vascular Endothelial Cells Is Modulated by a Coronary Artery Disease-Associated Genetic Variant and Influences Monocyte Transendothelial Migration. *J. Am. Heart Assoc.* **9**, e014333 (2020).

112. Lei, R. X. *et al.* Influence of a single nucleotide polymorphism in the P1 promoter of the furin gene on transcription activity and hepatitis B virus infection. *Hepatology* **50**, 763–771 (2009).

113. Ellinghaus, D. *et al.* Genomewide Association Study of Severe COVID-19 with Respiratory Failure. *N. Engl. J. Med.* (2020). doi:10.1056/NEJMoa2020283

114. Naitza, S. *et al.* A genome-wide association scan on the levels of markers of inflammation in sardinians reveals associations that underpin its complex regulation. *PLoS Genet.* **8**, (2012).

115. Pazoki, R. Cardiovascular disease, ABO locus, and markers of platelet functionality. *International journal of cardiology* **286**, 162–163 (2019).

116. Dai, X. ABO blood group predisposes to COVID-19 severity and cardiovascular diseases. *Eur. J. Prev. Cardiol.* **27**, 1436–1437 (2020).

117. Shokri, P. *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**,

118. Posch-Pertl, L. *et al.* Role of the ABO blood groups as a risk factor for retinal vein occlusion. *Ophthalmic Res.* (2022). doi:10.1159/000526874

119. Larsen, T. B. *et al.* ABO blood groups and risk of venous thromboembolism during pregnancy and the puerperium. A population-based, nested case-control study. *J. Thromb. Haemost.* **3**, 300–304 (2005).

120. Bawazir, W. M. Systematic Review and Meta-Analysis of the Susceptibility of ABO Blood Groups to Venous Thromboembolism in Individuals with Factor V Leiden. *Diagnostics (Basel, Switzerland)* **12**, (2022).

121. Shelton, J. F. *et al.* Trans-ancestry analysis reveals genetic and nongenetic associations with COVID-19 susceptibility and severity. *Nat. Genet.* **53**, 801–808 (2021).

122. Pairo-Castineira, E. *et al.* Genetic mechanisms of critical illness in COVID-19. *Nature* **591**, 92–98 (2021).

123. Zhang, Q. *et al.* Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science (80-.).* **370**, (2020).

124. Suhre, K. *et al.* A genome-wide association study of metabolic traits in human urine. *Nat. Genet.* **43**, 565–569 (2011).

125. Bröer, A. *et al.* Sodium translocation by the iminoglycinuria associated imino transporter (SLC6A20). *Mol. Membr. Biol.* **26**, 333–346 (2009).

126. Li, C. *et al.* Chemokine receptor CCR9 suppresses the differentiation of CD4(+)CD8αα(+) intraepithelial T cells in the gut. *Mucosal Immunol.* **15**, 882–895 (2022).

127. Passam, A. M. *et al.* Polymorphisms of Cx(3)CR1 and CXCR6 receptors in relation to HAART therapy of HIV type 1 patients. *AIDS Res. Hum. Retroviruses* **23**, 1026–1032 (2007).

128. Picton, A. C. P., Paximadis, M., Chaisson, R. E., Martinson, N. A. & Tiemessen, C. T. CXCR6 gene characterization in two ethnically distinct South African populations and association with viraemic disease control in HIV-1-infected black South African individuals. *Clin. Immunol.* **180**, 69–79 (2017).

129. Haniffa, M. *et al.* Human tissues contain CD141hi cross-presenting dendritic cells with functional homology to mouse CD103+ nonlymphoid dendritic cells. *Immunity* **37**, 60–73 (2012).

130. Saridaki, T. *et al.* FYCO1 mediates clearance of α-synuclein aggregates through a Rab7-dependent mechanism. *J. Neurochem.* **146**, 474–492 (2018).

131. Rothwell, S., Lilleker, J. B. & Lamb, J. A. Genetics in inclusion body myositis. *Curr. Opin. Rheumatol.* **29**, 639–644 (2017).

132. Downes, D. J. *et al.* Identification of LZTFL1 as a candidate effector gene at a COVID-19 risk locus. *Nat. Genet.* **53**, 1606–1615 (2021).

133. Rescenko, R. *et al.* Replication of LZTFL1 Gene Region as a Susceptibility Locus for COVID-19 in Latvian Population. *Virologica Sinica* **36**, 1241–1244 (2021).

134. Angulo-Aguado, M. *et al.* Association Between the LZTFL1 rs11385942 Polymorphism and COVID-19 Severity in Colombian Population. *Front. Med.* **9**, 910098 (2022).

135. Ishak, A. *et al.* The association of COVID-19 severity and susceptibility and genetic risk factors: A systematic review of the literature. *Gene* **836**, 146674 (2022).

136. Dobrijevic, Z. *et al.* The association of ACE1, ACE2, TMPRSS2, IFITM3 and VDR polymorphisms with COVID-19 severity: A systematic review and meta-analysis. *EXCLI J.* **21**, 818–839 (2022).

137. Saengsiwaritt, W., Jittikoon, J., Chaikledkaew, U. & Udomsinprasert, W. Genetic polymorphisms of ACE1, ACE2, and TMPRSS2 associated with COVID-19 severity: A systematic review with meta-analysis. *Rev. Med. Virol.* **32**, e2323 (2022).

138. Yildirim, Z., Sahin, O. S., Yazar, S. & Bozok Cetintas, V. Genetic and epigenetic factors associated with increased severity of COVID-19. *Cell Biol. Int.* **45**, 1158–1174 (2021).

139. Deb, P. *et al.* Association of HLA gene polymorphism with susceptibility, severity, and mortality of COVID-19: A systematic review. *HLA* **99**, 281–312 (2022).

140. Castelli, E. C. *et al.* MUC22, HLA-A, and HLA-DOB variants and COVID-19 in resilient super-agers from Brazil. *Front. Immunol.* **13**, 975918 (2022).

ANEXO 1

DO INÍCIO DO DOUTORADO ATÉ A CONCLUSÃO DESTA TESE

Meu ingresso no doutorado foi graças a um projeto bem diferente do que apresentei aqui: Havia redigido um projeto envolvendo genética do esporte, como uma tentativa de trabalharmos com um fenótipo comportamental que não fosse doença, sendo uma oportunidade ótima de trabalhar um estigma sobre a genética do comportamento (genética do comportamento = genética dos transtornos mentais). Alem disso, percebi a riqueza de fenótipos esportivos disponíveis e as possibilidades de se discutir seleção e adaptação. Dei início a revisão de literatura e a me inteirar sobre o tema. Li muita coisa, participei de reuniões na faculdade de educação física e terapia ocupacional da UFMG, no laboratório de pesquisas em esportes de combate (LAPEC), o qual meu então coorientador de doutorado, professor Maicon Rodrigues de Albuquerque, coordena. Professor Maicon foi muito solicito em nos receber quando conversamos sobre o projeto. Aprendi muito nas reuniões em seu laboratório sobre o fenótipo (minha maior limitação), e 1 ano se passou enquanto eu cumpria meus créditos e estudava o fenótipo.

Em 2020 o mundo parou: a organização mundial de saúde havia decretado crise sanitária em março daquele ano. Um novo coronavirus havia surgido na China e começado a se espalhar. Comércios, aeroportos e grandes eventos cancelados. Dentre estes grandes eventos, as olimpíadas no Japão.

Os centros de treinamentos e eventos cancelados, e isso prejudicou nossa coleta de fenótipo, que estava programada para começar em abril/maio daquele ano. A interação com os atletas não seria possível, e depois de um tempo com eles todos parados a mensuração do fenótipo também estaria comprometida. O LBI passou a fazer parte do enfrentamento da pandemia no auxílio do diagnóstico por PCR. As pesquisas no ICB foram paralisadas e apenas os Laboratórios parceiros do diagnóstico estavam autorizados a continuar funcionando. Os pós-graduandos foram convidados a ajudar nesse processo e suas pesquisas pessoais foram suspensas temporariamente. Com o tempo, os laboratórios conseguiram bolsas para contratar profissionais como técnicos para atuar no diagnóstico, e os pós-graduandos foram liberados para trabalhar na pesquisa com COVID-19. Os professores nos ofereceram projetos para conduzirmos e nos deram espaço para pensar em algumas ideias. Um dos biomarcadores genéticos que eu pensava em trabalhar no meu projeto com

esporte e já havia comprado os insumos apresentava grande plausibilidade biológica para o prognóstico da COVID-19: a deleção de 287 pares de base no gene *ACE1*. No esporte, estudada por sua influência na performance aeróbica e anaeróbica. Na COVID-19, sua interação com o gene *ACE2*, codificador da principal proteína de entrada celular para o virus SARS-CoV-2 chamou a atenção em primeiros estudos populacionais de prognóstico e suscetibilidade, encontrando resultados interessantes. Tínhamos acesso para pesquisa o restante das amostras utilizadas no diagnóstico de pacientes que estavam sendo internados no hospital Eduardo de Menezes, e apresentei a ideia de genotiparmos estes pacientes para esta deleção. Professor Renato acrescentou a possibilidade de dosarmos a expressão dos genes *ACE2* e *TMPRSS2*, fundamentais para a entrada do virus nas células humanas, nessa amostra. Um grupo no Rio de Janeiro com o qual o professor Renato (hoje meu coorientador) possuía parceria já estava conduzindo algo semelhante e começamos a trabalhar juntos em um projeto com a expressão destes genes, e comigo trabalhando em paralelo com a genotipagem da deleção no *ACE1.* Foi uma ótima oportunidade para aprender sobre ensaios de expressão gênica, e a analisar os tipos de dados gerados. Logo saiu nossa primeira publicação em genética humana e COVID-19. Um trabalho publicado na Scientific Report (doi: 10.1038/s41598-021-88944-8) onde dosamos a expressão dos genes *ACE2* e *TMPRSS2* em nossa amostra em Belo Horizonte e na amostra do Rio de Janeiro, e no qual apresentamos resultados interessantes sobre o envolvimento do *ACE2* no prognóstico e a razão *TMPRSS2*/*ACE2* sendo associada ao risco de maior dificuldade respiratória. Comparamos também a expressão destes genes em amostras de tecido nasofaríngeo e de tecido bronco alveolar, onde não encontramos diferenças.

Participei ainda em 2020 de um grande projeto de monitoramento conduzido na cidade de Betim-MG, onde monitoramos os casos de COVID-19 utilizando uma amostra com 3239 voluntários devidamente selecionados pela idade, sexo e região geográfica. Fizemos um estudo temporal envolvendo 3 grandes etapas de coleta, e este estudo envolveu inúmeros profissionais de saúde. O trabalho foi publicado na Frontiers in Microbiology (doi: 10.3389/fmicb.2022.799713).

Depois de um tempo começaram a aparecer variantes de interesse e preocupação do virus SARS-CoV-2, as chamadas VOI e VOC. Dessa forma, viu-se a necessidade de se iniciar um processo de vigilância genômica com maior rigorosidade para monitorar a dispersão destas variantes. Identificamos em Minas Gerais a variante

Gamma (P1) em fevereiro de 2021. Demos início ao monitoramento desta variante no estado de Minas. Nessa época, havíamos protocolado um método de identificação de variantes rápido e prático, que poderia substituir a necessidade de protocolos de sequenciamento na grande maioria das amostras, o que foi fundamental para a velocidade da reportagem de dados de vigilância para as autoridades, já que conseguíamos identificar de forma mais rápida um número bem maior de amostras do que em um processo de sequenciamento NGS (New Generation Sequencing). Esse protocolo foi baseado em genotipagem de polimorfismos de nucleotídeo único no genoma do virus que os diferenciavam de demais variantes virais (dx.doi.org/10.17504/protocols.io.buf2ntqe) e mantivemos o sequenciamento para amostras que apresentavam problemas na identificação por este protocolo. Monitoramos desta forma o avanço da variante Gamma em Minas Gerais. Participei ativamente do início do estabelecimento deste protocolo no LBI, e o resultado do trabalho com a variante Gamma foi publicado na Viruses (doi: 10.3390/v14122747).

O processo foi passado para demais alunos do LBI e o monitoramento das variantes do SARS-CoV-2 permanece em atividade até hoje. Passei então a focar em meus projetos com genética humana. Iniciei uma parceria com uma aluna de mestrado do laboratório (Isabela Braga) para investigar a influência dos marcadores que já estávamos trabalhando no LBI com a perda de olfato e de paladar, sintomas que foram comumente relatados nos pacientes acometidos pela COVID-19, em especial aos pacientes com quadros clínicos ditos leves. O trabalho foi publicado na Frontiers in cellular and infection microbiology (doi: 10.3389/fcimb.2022.905757).

Em paralelo aos meus experimentos, conduzi uma revisão sistemática com o mestre e doutorando Diego Menezes e a Dra. Paula Fonseca, pós doutoranda no LBI, envolvendo uma análise cientométrica dos estudos de vigilância genômica publicados no Brasil. O trabalho organizou todos os estudos de vigilância publicados com amostras brasileiras até setembro de 2021, e foi publicado no periódico Viruses (doi: 10.3390/v14122715). Com o tempo percebi que deveria transformar todo o trabalho executado durante esse momento em minha tese. A pandemia não parecia ter fim certo, e eu precisava me qualificar. Renunciar ao projeto que me abriu portas ao doutorado era necessário, apesar de manter um desejo de colaborar futuramente com ele.

Durante toda a pandemia trabalhamos muito, com todas as dificuldades, e produzimos muito conhecimento. Tive a oportunidade de aumentar meu

conhecimento em estatística, e em redação de texto. Colaborei, publiquei, escrevi muito, tive ideias. Muitas puderam ser conduzidas, tantas outras não. Mas conclui uma etapa importantíssima em minha vida. Mesmo com todas as mudanças de planos em minha trajetória, algo se manteve: meu amor por ciência, genética e evolução, principalmente humana. Cresci muito durante todo o processo, aprendi a lidar com adversidades, tive contato mesmo que indireto com a história dos pacientes. Saber que muito dos dados de falecimento com os quais lidei eram de pacientes que executei o diagnóstico no início quando ainda estavam vivos é marcante, é triste, emociona. Saber que esta tese é resultado de toda minha trajetória, não apenas na epidemiologia, mas resultado de todos os laboratórios nos quais já trabalhei, também emociona. Tudo fez e faz parte do profissional que agora se forma doutor, e jamais me esquecerei no caminho que percorri até este momento. Apenas agradeço a todos os professores que passaram por minha vida, desde minha formação básica até a graduação e pós-graduação. Sem vocês, eu não seria quem sou.