UNIVERSIDADE DO ESTADO DO PARÁ
INSTITUTO EVANDRO CHAGAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA PARASITÁRIA NA AMAZÔNIA

ANÁLISE DE POLIMORFISMOS DOS GENES *ACE-2* E *TMPRSS2* EM PROFISSIONAIS QUE ATUARAM CONTRA A COVID-19 EM BELÉM- PA

CAROLINY SOARES SILVA

Belém – Pará
2025

CAROLINY SOARES SILVA

ANÁLISE DE POLIMORFISMOS DOS GENES *ACE-2* E *TMPRSS2* EM PROFISSIONAIS QUE ATUARAM CONTRA A COVID-19 EM BELÉM- PA

Tese apresentada ao Programa de Pós-Graduação em Biologia Parasitária na Amazônia da Universidade do Estado do Pará/Instituto Evandro Chagas como requisito para a obtenção do grau de Doutor em Biologia Parasitária da Amazônia.

Orientador (a): Profa Dra. Luana Nepomuceno Gondim Costa Lima

Belém – Pará
2025

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da UEPA / SIBIUEPA

Silva, Caroliny Soares

Análise de polimorfismos dos genes ACE-2 e TMPRSS2 em profissionais que atuaram contra a covid-19 em Belém-Pa / Caroliny Soares Silva. – Belém: UEPA, 2025.

140f.: il.

Orientadora: Prof.ª Dra. Luana Nepomuceno Gondim Costa Lima

Tese (Doutorado) – Universidade do Estado do Pará, Centro de Ciências Biológicas e da Saúde, Programa de Pós-graduação em Biologia Parasitária na Amazônia, Belém, 2025.

1. COVID-19. 2. Polimorfismo genético. 3. ACE-2. 4. TMPRSS. I. Universidade do Estado do Pará. II. Título.

CDD 23.ed. 616.24144

Elaborada por Josicléia Garcia Vieira - CRB-2 /562

CAROLINY SOARES SILVA

ANÁLISE DE POLIMORFISMOS DOS GENES *ACE-2* E *TMPRSS2* EM PROFISSIONAIS QUE ATUARAM CONTRA A COVID-19 EM BELÉM- PA

Tese apresentada ao Programa de Pós-Graduação em Biologia Parasitária na Amazônia da Universidade do Estado do Pará/Instituto Evandro Chagas como requisito para a obtenção do grau de Doutor em Biologia Parasitária da Amazônia.

Orientadora: Profa. Dra. Luana Nepomuceno Gondim Costa Lima

Banca Examinadora:

Prof. Dr. Carlos Alberto Marques de Carvalho
Universidade do Estado do Pará

Prof. Dra. Jannifer Oliveira Chiang
Instituto Evandro Chagas

Prof. Dr. Jorge Rodrigues de Sousa
Universidade do Estado do Pará

Prof. Dr. Ricardo José de Paula Souza e Guimarães
Seção de Epidemiologia/IEC/SVSA/MS

Belém, 28 de outubro de 2025

“Na luta do bem contra o mal, é sempre o povo que morre. ”
Eduardo Galeano

Dedico à minha avó e à Solange.
A saudade e os agradecimentos são eternos.

AGRADECIMENTOS

Primeiramente, gostaria de agradecer à Deus, por toda bondade, proteção e ajuda que me foi concedida.

À minha avó Ana, por todo amor e apoio. Por todos os dias das mães e pais. Por todas as idas para me buscar na escola. Por toda felicidade em cada conquista. Obrigada por me ajudar a ser a pessoa que sou hoje. Espero que esteja feliz e com orgulho.

Aos meus pais, Dora Soares e Rogério, por todo o amor, carinho, compreensão e apoio diário. Obrigada por todo o apoio dado, mesmo nos momentos mais difíceis. Não existem palavras que expressem toda a gratidão que sinto.

Ao Gabriel, meu melhor amigo e irmão. Obrigada por ter me apoiado em todos os momentos e por nunca me deixar só. Te amo. Espero um dia ter metade da sabedoria que você tem.

Agradeço à minha orientadora Luana Nepomuceno pela oportunidade e ensinamentos.

Agradeço à banca por pela disponibilidade e considerações sobre esse trabalho.

Agradeço à professora Cristiane Cunha por toda a ajuda, pelos conselhos e por todos os ensinamentos nesses dez anos. Sempre serei grata pela oportunidade de trabalhar com a senhora.

À Fundação Amazônia de Amparo a Estudos e Pesquisas (FAPESPA) por ter me oferecido uma bolsa no Programa de Pós-graduação, que ajudou durante a minha formação.

À todos do Instituto Evandro Chagas, pela colaboração, em especial ao Marcos Jessé, por toda a parceria nos trabalhos publicados e pela amizade. Obrigada pela ajuda e ensinamentos que foram essenciais para o meu trabalho.

Aos meus colegas do Laboratório de Micobactérias e Plataforma de Bioensaios do Instituto Nacional de Infectologia da Fundação Oswaldo Cruz por todo o apoio dado para terminar esse trabalho.

Ao Luis Felipe por todo amor e companheirismo. Obrigada por me ajudar a ser mais forte para encarar os dias e a não desistir. Amo você e a Elis.

E assim como no TCC e dissertação, dedico e agradeço ao Lupin por todo amor e por estar sempre perto de mim, sendo um grande companheiro desde 2014.

RESUMO

A covid-19 é uma doença infectocontagiosa causada pelo vírus SARS-CoV-2, que pode ser sintomática e com desfechos desfavoráveis, principalmente, pulmonares. A evolução clínica dela pode ser fruto de diversos fatores, tais como os ambientais, imunológicos e genéticos. Nesse sentido, o período pandêmico da covid-19 trouxe danos incalculáveis para a humanidade e a pandemia assolou na sua primeira onda, sobretudo, os profissionais da estrutura clínica-hospitalar de todo o mundo. O objetivo deste trabalho foi analisar o perfil genotípico e alélico de SNPs dos genes *ACE-2* e *TMPRSS2* que poderiam atuar como fator de risco para covid-19 em profissionais que trabalharam em instituições de saúde durante a primeira onda da pandemia. Trata-se de um estudo de coorte retrospectivo, com amostras coletadas de profissionais das categorias de serviços gerais, administrativa e da saúde de 10 instituições de saúde de Belém, Pará, com um total de 123 pacientes sintomáticos e 91 assintomáticos. Foi avaliada a sintomatologia dos indivíduos após a análise do questionário epidemiológico e em indivíduos sintomáticos, houve a realização da tomografia do tórax para a avaliação do comprometimento pulmonar. As frequências de SNP foram obtidas por meio de sequenciamento parcial de Sanger. O genótipo C/C do SNP *rs961360700* do gene foi encontrado em todas as amostras analisadas. Para o SNP *rs2298659*, 68,24% do genótipo C/C e 29,8% do genótipo C/T foram detectados, enquanto apenas 1,29% do genótipo T/T foi detectado. A presença desse polimorfismo foi significativa na alteração dos fenótipos sintomáticos e a correlação entre indivíduos sem comorbidades prévias e o início dos sintomas. Esses achados destacam a importância de investigar esses polimorfismos em outras populações, bem como analisar outros SNPs nesses genes responsáveis pela entrada viral, o que pode ajudar a entender o agravamento dos casos de covid-19.

Palavras-chave: Profissional da Saúde; Enzima de Conversão de Angiotensina 2; covid-19; SARS-CoV-2.

ABSTRACT

COVID-19 is an infectious disease caused by the SARS-CoV-2 virus, which may present with symptoms and lead to unfavorable outcomes, particularly pulmonary complications. Its clinical course can be influenced by multiple factors, including environmental, immunological and genetic determinants. During the first wave of the pandemic, COVID-19 caused incalculable damage to humanity and had a particularly devastating impact on healthcare professionals worldwide. The objective of this study was to analyze the genotypic and allelic profiles of SNPs in the ACE-2 and TMPRSS2 genes that could act as a risk factor for COVID-19 among professionals working in health institutions during the first pandemic wave. This retrospective cohort study included samples collected from professionals in general, administrative and healthcare services categories from 10 health institutions in Belém, Pará. The cohort comprised 123 symptomatic and 91 asymptomatic individuals. Clinical symptoms were assessed through an epidemiological questionnaire and symptomatic individuals underwent chest tomography to evaluate pulmonary involvement. These data were correlated with the frequencies of SNPs rs2298659 in the TMPRSS2 gene and rs961360700 in the ACE-2 gene. SNP frequencies were determined through partial Sanger sequencing. The C/C genotype of SNP *rs961360700* in ACE-2 was observed in all analyzed samples. For SNP *rs2298659*, 68.24% of individuals carried the C/C genotype, 29.8% the C/T genotype, and only 1.29% the T/T genotype. The presence of this polymorphism showed a significant association with symptomatic phenotypes and the onset of symptoms in individual without previous comorbidities. These findings underscore the importance of investigating these polymorphisms in other populations, as well as analyzing additional SNPs in these genes involved in viral entry, which may contribute to understanding the progression and severity of COVID-19 cases.

Keywords: Healthcare Professional; Angiotensin Converting Enzyme 2; COVID-19; SARS-CoV-2.

SUMÁRIO

LISTA DE FIGURAS

LISTA DE QUADROS

LISTA DE SIGLAS E ABREVIAÇÕES

Angiotensinogênio angiotensina I	Ang I
Angiotensinogênio angiotensina II	Ang II
Assintomáticos	AS
Células apresentadoras de antígeno	APC
Coronavírus disease 2019	COVID-19
Corovavírus	CoVs
Domínio fusão citoplasmático	CP
Domínio heptad repeat 1	HR1
Domínio heptad repeat 2	HR2
Domínio transmembrana	TM
Gene relacionado aos ETS	ERG
Envelope	E
Enzima conversora de angiotensina	ECA
Enzima conversora de angiotensina 2	ACE-2
Grupo de Estudo de Coronavírus	CSG
Imunoglobulina G	IgG
Imunoglobulina M	IgM
Membrana	M
Nucleocapsídeo	N
Open Reading Frames	ORF
Organização Mundial de Saúde	OMS
Peptídeo de Fusão	PF
Peptídeo de fusão interno	PFI
Polimorfismo de nucleotídeo único	SNP
Poliproteínas pirofosfatase –1a	Ppa1a
Poliproteínas pirofosfatase –1b	PPa1B
Protease do tipo tripsina das vias aéreas humanas	HAT
Protease transmembranar serina 2	TMPRSS2
Proteínas Não Estruturais	PNE
Reação em cadeia da polimerase em tempo real da transcriptase reversa	RT-qPCR
Receptores de reconhecimento de padrões	PRR
Receptores NOD-like	NLRs
Receptores Toll-like	TLRs
Segundo sítio proteolítico	S2′
Severe Acute Respiratory Syndrome Coronavírus 2	SARS-CoV-2
Síndrome Respiratória do Oriente Médio	MERS
Sintomáticos	SI
Sintomáticos sem comprometimento pulmonar	SSP
Sistema renina-angiotensina-aldosterona	SRAA
Spyke	S
Termo de consentimento livre esclarecido	TCLE

1 INTRODUÇÃO

A doença causada pelo coronavírus 2019 (covid-19 – *Coronavirus disease 2019*) é causada pelo coronavírus 2, agente da síndrome respiratória aguda grave (SARS-CoV-2 – *Severe Acute Respiratory Syndrome Coronavirus 2*), um vírus descoberto no final de dezembro de 2019, que apresentou uma alta taxa de infectividade (SENEVIRATHNE *et al*., 2024). O patógeno afeta principalmente o sistema respiratório, causando resfriados, embora outros órgãos possam ser acometidos. Os sintomas mais comuns do COVID -19 incluem febre, tosse, falta de ar e dor muscular (LIPPI; SANCHIS-GOMAR; HENRY, 2023)

O vírus se espalha principalmente através de gotículas respiratórias quando os indivíduos falam, tossem, espirram ou por meio de secreções nas orofaríngeas e salivares (SHARMA; AHMAD FAROUK; LAL, 2021). A infecção também pode ocorrer pelo contato com superfícies contaminadas e posterior contato com a face. Se o indivíduo tocar em uma superfície contaminada com partículas virais e levar para a membranas mucosas da boca, olhos e nariz ele pode se infectar com o SARS-CoV-2. O vírus também foi encontrado em sangue e fezes, podendo indicar a possibilidade de múltiplas rotas de transmissão (CHUNG *et al*., 2024a; ŠUŠAK *et al*., 2024). Indivíduos assintomáticos também conseguem transmitir amplamente o SARS-CoV-2, espalhando aerossóis infecciosos contendo vírions encapsulados produzidos após fala, canto, tosse ou espirro (TELLIER, 2022).

O SARS-CoV-2 é um vírus de RNA que apresenta uma alta taxa de erro durante a sua replicação. Esses erros podem causar constantes mutações genéticas, resultando na criação de novas variantes. Uma das mutações mais detectadas são as mutações do tipo *missense* na proteína *spike*, que podem aumentar a sua transmissibilidade e virulência, ocasionando a redução da eficácia da vacina (HADJ HASSINE, 2022). Com o surgimento de novas variantes, pacientes que já foram infectados podem sofrer uma reinfecção devido à capacidade de escape imunológico. A taxa de reinfecção era de cerca 0,94% até a chegada da variante Ômicron, onde foi detectado um aumento na taxa para cerca de 4,1% (CHEN *et al*., 2024)

Alguns outros fatores podem influenciar na proteção ou no desenvolvimento da covid-19, como dieta, nutrição, vacinação, fatores imunológicos e genéticos. O aumento de citocinas anti-inflamatórias e a diminuição da enzima conversora de

angiotensina 2 (*Angiotensin-converting enzyme 2* - ACE-2) de uma forma ativa no sistema cardiovascular estão associados à inflamação pulmonar, fibrose e danos progressivos ao tecido (ASHOUR, 2022.). Estudos mostram que o SARS-CoV-2 tem como o principal alvo o sistema cardiovascular, onde as doenças cardíacas foram as comorbidades mais encontradas em pacientes em todas as fases de infecção por covid-19. Uma inflamação sistêmica pode ocorrer em indivíduos com o ACE-2 e TMPRSS2 comprometidos pelo SARS-CoV-2, causando um aumento nos marcadores inflamatórios como o IL-6, IL-2 e TNF-α contribuindo ainda mais para a formação de uma lesão cardíaca (BÖHM *et al*., 2020). A partir disso, a indução de anticorpos e a ativação de células T através da vacinação podem ajudar a evitar a infecção e o agravamento da doença (ZHANG *et al*., 2023).

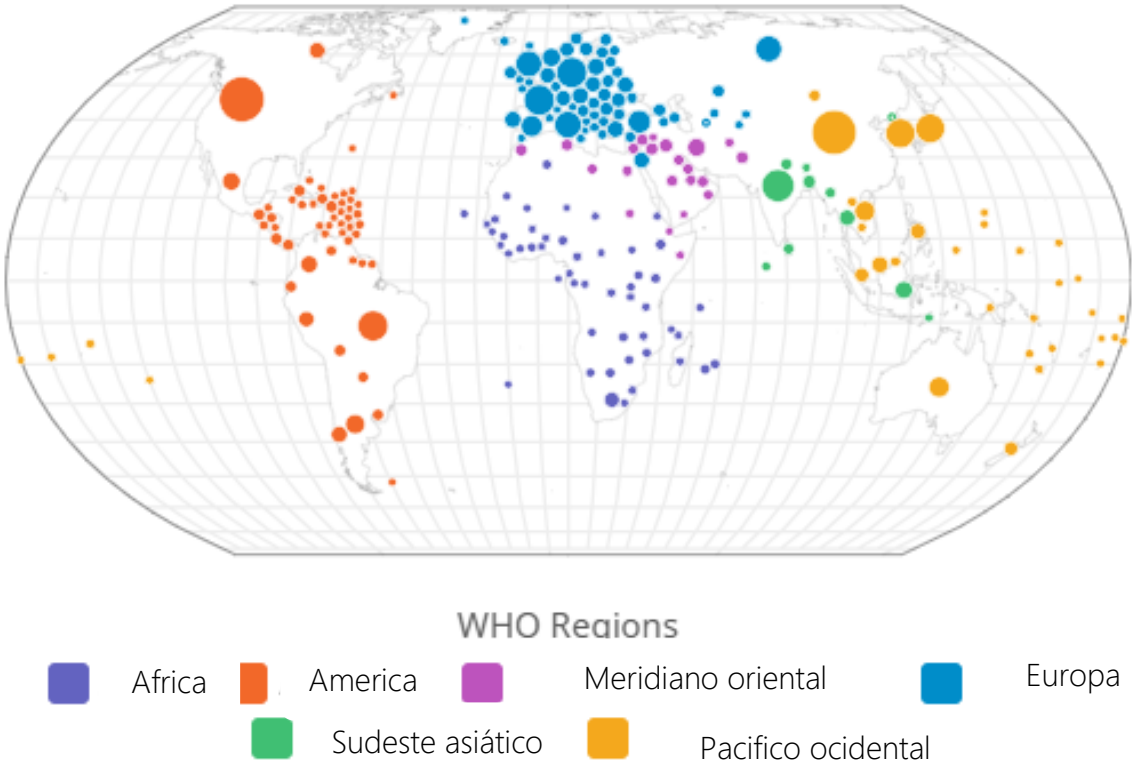
2 REFERENCIAL TEÓRICO

2.1 EPIDEMIOLOGIA E HISTÓRICO

Em dezembro de 2019, a Organização Mundial da Saúde (OMS) foi informada sobre casos de pneumonia de origem desconhecida na cidade de Wuhan, província de Hubei e, rapidamente se espalhou pelo mundo (LUNDBERG *et al*., 2024). Posteriormente, o agente etiológico foi identificado e chamado de 2019-nCoV. Em fevereiro de 2020, a OMS oficialmente nomeou a doença de coronavírus 2019 (covid-19) e posteriormente Grupo de Estudo de Coronavírus (CSG) do Comitê Internacional de Taxonomia de Vírus propôs nomear o novo coronavírus como SARS-CoV-2 (WHO, 2025).

Atualmente, no mundo existem mais de 700 milhões de casos confirmados da doença, no Brasil já são 37,7 milhões de casos confirmados e 715.108 óbitos acumulados (FIGURA 1) (WHO, 2025).

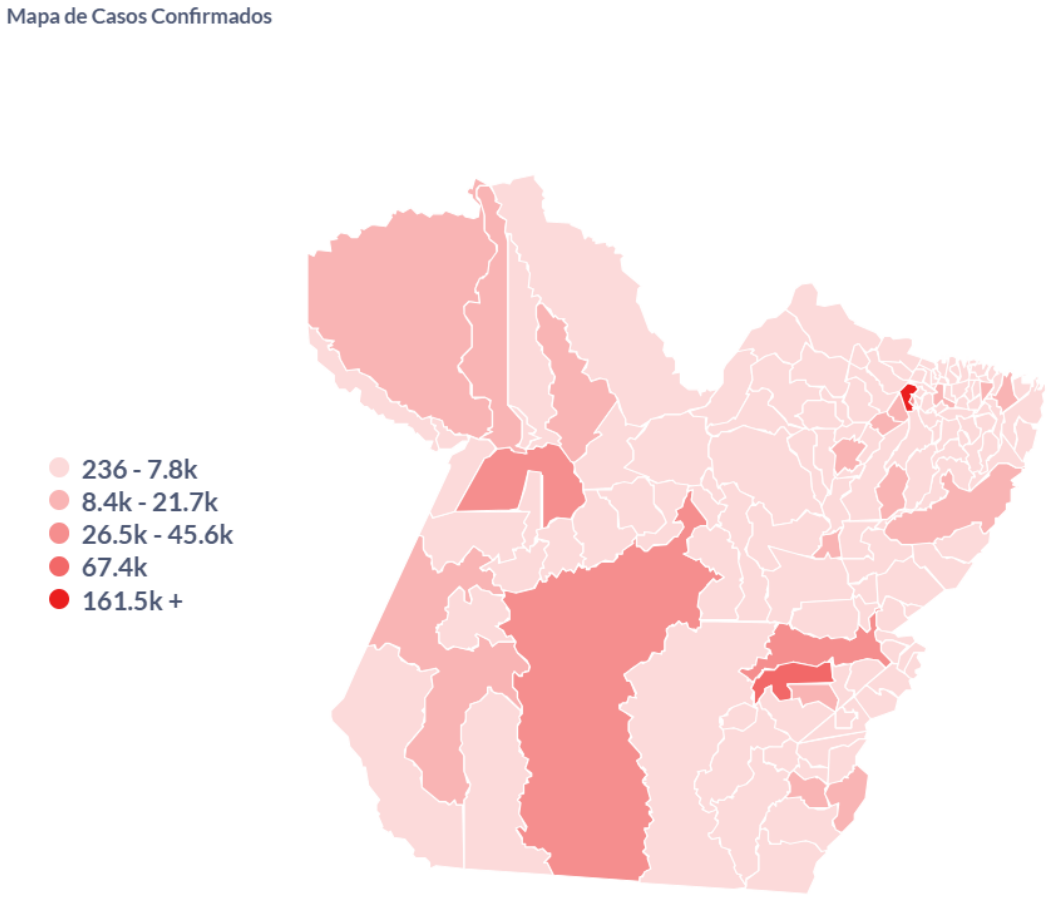
Figura 1 - Distribuição de casos confirmados de covid-19 no mundo



Fonte: WHO, 2025. Dados epidemiológicos do acumulado de casos no mundo desde o início da contagem oficial pela OMS. Acumulados de casos nos Países: Estados Unidos: 103.436,829; China: 99.381.761; Brasil: 37.797.516; Índia: 45.056.073.

Na região Norte do País, foram confirmados aproximadamente 3 milhões de casos de covid-19, apresentando a segunda maior taxa de mortalidade entre as regiões, com 283,16/100 mil hab. O estado do Pará apresenta um acumulado de casos com um total de 905.316, com um total de 19.329 óbitos confirmados, com uma taxa de mortalidade próxima à da região Norte com 224,68/100 mil hab. (FIGURA 2) (BRASIL, 2025).

Figura 2 – Mapa da incidência acumulativa de casos confirmados de covid-19 no estado do Pará até 2025



Fonte: PARÁ, 2025. Dados epidemiológicos do acumulado de casos no estado do Pará desde o início da contagem oficial pela Secretaria de Saúde do Estado do Pará e pelo Ministério da Saúde. Municípios com maiores incidências: Belém do Pará: 161.692 hab.; Parauapebas: 67.438 hab.; Santarém: 45.867 hab.; Marabá: 28.206 hab.; Altamira: 26.589 hab.

2.2 DIAGNÓSTICO

A detecção do SARS-CoV-2 teve que sofrer diversas mudanças devido ao surgimento de novas variantes. O desenvolvimento de novas técnicas visa analisar os ácidos nucleicos virais, como os testes moleculares; analisar a presença de antígenos ou a presença de anticorpos contra o SARS-CoV-2, como o teste de antígeno e sorológico, respectivamente (FERNANDES *et al*., 2022).

Para a detecção molecular dos ácidos nucleicos, são coletadas amostras de swab nasofaríngeo, apresentando uma grande sensibilidade para a detecção viral. As coletas são analisadas através da reação em cadeia da polimerase em tempo real da transcriptase reversa (RT-qPCR), sendo necessário serem realizadas entre o dia 0 (início) e dia 4 do início dos sintomas após a infecção, período em que a carga viral é mais alta (Alsharif; Qurashi, 2021). O ensaio de RT-qPCR foi desenvolvido para detectar diferentes genes, como o gene do nucleocapsídeo, gene do envelope, gene spike e as regiões ORF1b ou ORF8 do genoma do SARS-CoV -2 (RAi *et al.*, 2021).

Os testes sorológicos detectam as células do sistema imunológico, como imunoglobulinas e anticorpos, inclusive a imunoglobulina M (IgM) e a imunoglobulina G (IgG), moléculas produzidas após o sistema imunológico detectar a presença viral (YÜCE; FILIZTEKIN; ÖZKAYA, 2021). A presença de anticorpos IgM pode indicar a recente infecção viral, onde em pacientes sintomáticos a duração média de anticorpos é de 5 dias, enquanto a IgG pode ser detectada posteriormente, em média de 14 dias, indicando uma infecção posterior. Em alguns casos, pode ser detectada a persistência do IgM anti-SARS-CoV-2 por até 8 meses após a infecção (BICHARA *et al*., 2021).

2.3 SINTOMATOLOGIA E QUADROS CLÍNICOS

Os sintomas mais comuns da covid-19 são febre, tosse e dispneia, e em alguns casos existe o relato de diarreia. Pessoas com idade acima de 65 anos, apresentam um risco maior de desenvolverem a forma mais grave da doença devido a presença de comorbidades que podem agravar os casos. Entretanto, adultos jovens também podem ser hospitalizados por agravamento dos quadros clínicos, mesmo com a ausência de comorbidades pré-existentes (OCHANI *et al*., 2021).

A infecção e sintomatologia podem ser divididas em três estágios sendo que o primeiro é caracterizado como período de incubação após a infecção por SARS-CoV-2, onde ocorre a replicação viral sem o aparecimento de sintomas aparentes. Posteriormente ocorre a disseminação, sendo diagnosticada pela pneumonia aguda e sintomas como febre e tosse. Caso o paciente não consiga se recuperar da segunda fase, ele entrará no estágio três, considerado o mais grave, podendo ter uma hiperinflamação sistêmica, lesão renal, lesão hepática e gastrointestinal *(*CHUNG *et al*., 2024a).

2.4 PREVENÇÃO DA DOENÇA E TRATAMENTO

A forma mais eficaz para prevenir a covid-19 é através da vacinação (CHUNG *et al*., 2024). As vacinas apresentam um papel importante na prevenção e transmissão, onde estas podem ser desenvolvidas a partir do patógeno ou de seus fragmentos, induzindo a resposta imune adaptativa, com a indução da produção de anticorpos e a ativação da imunidade celular (ZHAO *et al.*, 2024).

As vacinas tradicionais com patógeno inteiro são eficazes, assim como as vacinas que possuem os fragmentos antigênicos do patógeno, que além da sua capacidade de induzir os anticorpos, apresentam menos efeitos colaterais e são mais seguras para indivíduos que apresentam imunossupressão (VARTAK; SUCHECK, 2016). Para controlar a disseminação do vírus, diversos tipos de vacinas foram testadas, como as vacinas de proteína recombinante, vacinas de vetor de adenovírus, vacinas de DNA e vacinas de mRNA, sendo que esta última apresentou um papel importante no controle da pandemia (FANG *et al*., 2022).

Contudo, o SARS-CoV-2 constantemente sofre mutações formando novas variantes que conseguem causar novas infecções. Um exemplo disso é o surgimento da variante Ômicron JN.1, que não é detectado pelos anticorpos neutralizantes, diminuído assim a eficácia vacinal, sendo necessária constantes atualizações nas novas vacinas (YANG *et al*., 2024). Com isso, são necessárias novas formas que auxiliem no tratamento e na diminuição dos sintomas da covid-19 enquanto novas vacinas são fabricadas.

A maioria das estratégias utilizadas no tratamento foram baseadas nos tratamentos utilizados contra SARS e Síndrome respiratória do Oriente Médio (MERS), pois apresentam características semelhantes. Atualmente diversos medicamentos estão sendo estudados em ensaios clínicos e aprovados pelas agências de saúde (CHUNG *et al*., 2024). Os tratamentos utilizados no tratamento da doença podem ser divididos em dois grandes grupos de acordo com a sua ação: terapias para auxiliar o hospedeiro e a utilização de antivirais, como Baricitinibe, Paxlovid e Remdesivir (YUAN *et al*., 2023).

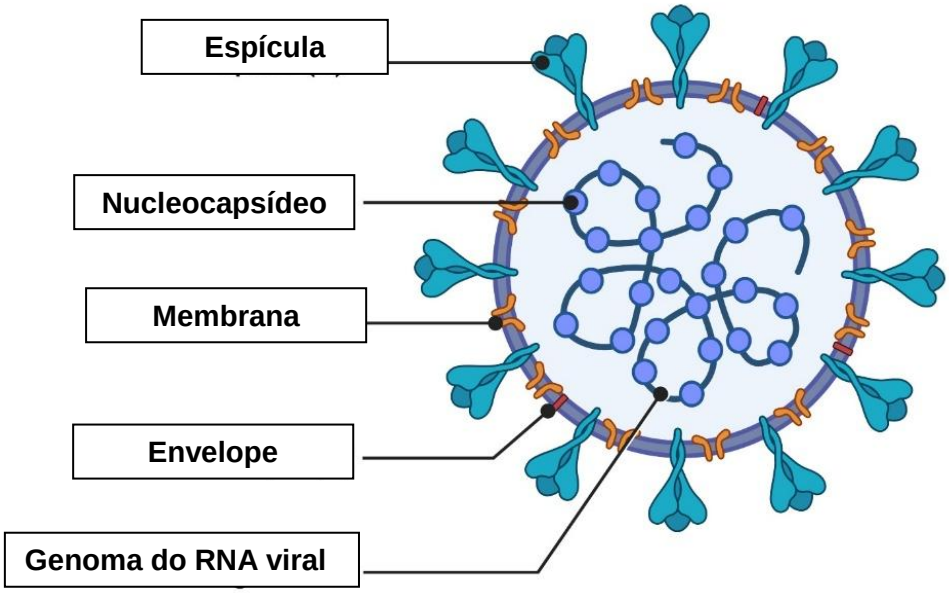
2.5 CORONAVÍRUS

2.5.1 Classificação, morfologia e genoma

O SARS-CoV-2 é um β-coronavírus envelopado esférico com RNA de sentido positivo com fita simples não segmentado e apresenta um genoma com tamanho de cerca de 29,9 kb (HADJ HASSINE, 2022). O vírus pode infectar o sistema nervoso central, respiratório e o gastrointestinal de humanos, outros mamíferos e aves e são responsáveis por causar diversos problemas de saúde que podem variar de infecções simples a infecções severas e persistentes que afetam principalmente as vias respiratórias, neurológicas e entéricas (TEBHA *et al*., 2024). Os coronavírus (CoVs) pertence a ordem *Nidovirales*, família *Coronoviridae*, subfamília *Orthocoronavinae*, sendo subdividida em quatro gêneros: *Alphacoronavirus, Betacoronavirus*, *Gammacoronavirus* e *Deltacoronavírus* (HUSSEIN *et al*., 2024).

Os CoVs possuem um contorno esférico, com o diâmetro do vírion variando de 60 a 140 nm e proteínas estruturais e não estruturais (PNE): *espícula* (S), *membrana* (M), *envelope* (E) e *nucleocapsídeo (N)* e proteínas acessórias (FIGURA 3). Apresentam espículas de 9 a 12 nm que se projetam na superfície viral, possuindo a aparência de uma coroa (MOHAMADIAN *et al*., 2021).

Figura 3 – Estrutura morfológica e principais componentes do coronavírus



Fonte: Hassine, (2021). Organização estrutural do coronavírus com distribuição das suas principais proteínas morfológicas e que estão envolvidas na entrada viral na célula do hospedeiro.

A proteína S possui a capacidade de induzir a resposta imune do hospedeiro, sendo dividida em duas subunidades: S1 e S2. A S1 auxilia na ligação ao receptor do hospedeiro, enquanto a S2 é responsável pela fusão viral na membrana do hospedeiro. A subunidade S2 é composta pelo peptídeo de fusão (PF), um segundo sítio proteolítico (S2), um peptídeo de fusão interno (PFI), domínio *heptad repeat* 1 (HR1), domínio *heptad repeat 2* (HR2), domínio transmembrana (TM) e domínio fusão citoplasmático (CP) (AISENBREY; BECHINGER, 2024).

A proteína M é a mais abundante encontrada no vírion, responsável pela definição do formato do envelope viral. Ela é responsável pela ligação ao nucleocapsídeo e possui uma alta diversidade no conteúdo de seus aminoácidos, contudo apresentam similaridade na sua estrutura em diferentes gêneros. A proteína E é um pequeno polipeptídeo que participa como canal iônico na membrana e possui função na liberação do vírus, montagem e patogênese (NAL et al., 2005). A proteína N apresenta diversas funções, entre elas está a formação dos complexos com o genoma viral, facilitação da interação da proteína M na contagem do vírion e ao aumento da eficiência da transcrição do vírus (XIANG *et al*., 2024).

Após a entrada desse vírus na célula, o mRNA está pronto para ser traduzido em proteína. O genoma desse vírus possui o comprimento inferior a 30 kb, onde existem cerca de 14 quadros de leitura aberta (*Open Reading frames* - ORF) que codificam diversas proteínas que desempenham um papel na sua sobrevivência e virulência, incluindo as proteínas S, E, M e N (ASTUTI; YSRAFIL, 2020; WIEDMANN *et al*., 2024).

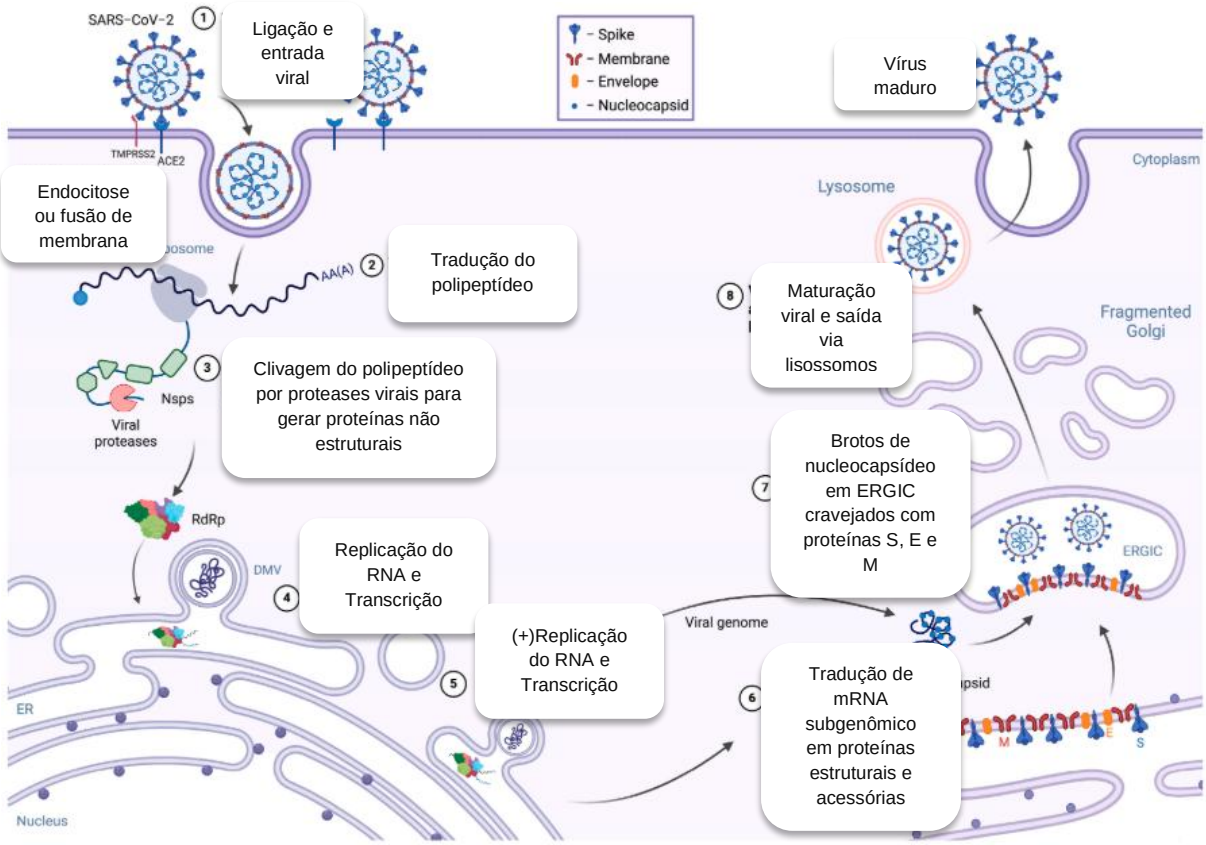
Posteriormente a entrada viral, as poliproteínas pirofosfatase –1a (Ppa1a) e a PPa1B são produzidas pelo quadro de leitura aberto 1a (ORF1a) e 1b (ORF1b), onde posteriormente dezesseis proteínas não estruturais são secretadas via proteólise Ppa1a e Ppa1b, onde estas sequestram a maquinaria de produção das células hospedeiras e utilizam para a replicação viral (VOSKO; ZIRLIK; BUGGER, 2023).

A primeira ORF apresenta cerca de 65% do genoma viral, entre elas estão a PNE3, PNE9, PNE10, PNE12, PNE15 e PNE16, que são responsáveis pela replicação viral (MOHAMANDIAN, *et al*., 2021).

2.5.2 Formas de entrada

A entrada do SARS-CoV-2 na célula acontece por meio de seu receptor à enzima conversora de angiotensina 2 (ACE-2), existente em vários órgãos, como os pulmões, rins, coração e trato gastrointestinal. A ligação ocorre pelo domínio de ligação da proteína S do SARS-CoV que se liga fortemente ao ACE-2 humano. As proteases das células alvo ativam a proteína S, clivando-a em subunidades S1 e S2. A S2, muda de conformação, incluindo a inserção de PF na membrana da célula alvo e exposição do domínio HR1. A interação entre os domínios de HR2 e HR1 formam feixes de seis hélices (6-HB), trazendo o envelope viral e a membrana celular em estreita proximidade para a fusão e entrada viral. As proteases humanas que clivam e ativam a proteína S no processo de fusão com a membrana humana, são a protease transmembranar serina 2 (TMPRSS2) e a protease do tipo tripsina das vias aéreas humanas (HAT) (KATIYAR *et a*l., 2024; LIMA; SOUSA; LIMA, 2020).

Figura 4 - Mecanismo de infecção por SARS-CoV-2 nas células do hospedeiro



Fonte: Katiyar *et al*., (2024). A entrada viral é facilitada pelo ACE-2 *e TMPRSS2.* Após a endocitose, o vírus utiliza a maquinaria celular para realizar a sua replicação e tradução multiplicando a quantidade de vírus no organismo, prontos para serem disseminados

O SARS-CoV-2 pode invadir as células por endossomos ou por fusão com a membrana plasmática. Além disso, é descrito um outro mecanismo de invasão viral, onde o vírus interage com as porfirinas. Em um estudo de acoplamento molecular, modelagem de homologia e análise de domínios conservados realizado por Wenzhong & Hualan (2020), verificou-se que as proteínas E, N e ORF3a apresentavam sítios de ligação ao ferro do heme, podendo a ORF3a dissociar o átomo de ferro e formar porfirinas. As proteínas E2 e E do SARS-CoV-2 por sua vez podem se ligar as porfirinas, sugerindo uma outra forma de invasão da célula hospedeira (LIMA; SOUSA; LIMA, 2020).

2.6 PATOGÊNESE E RESPOSTA IMUNOLÓGICA DO HOSPEDEIRO A INFECÇÃO PELO SARS-COV-2

A patogênese da covid-19 está diretamente correlacionada aos principais processos patológicos associados à inflamação ocasionando manifestações locais de inflamação, inflamação sistêmica aguda e inflamação sistêmica crônica de baixa intensidade. Além disso, existem os fatores envolvidos na virulência como o fácil reconhecimento do vírus por receptores celulares amplamente espalhados pelas células do hospedeiro e a supressão da resposta antiviral das células infectas e do sistema imunológico (GUSEV *et al*., 2022). A partir disso, a expansão viral, a morte celular e a entrada de produtos na corrente sanguínea, o amplo recrutamento de células imunológicas e ativação de sistemas de complemento levam à um inflamação sistêmica, que pode ser agravada de acordo com as características prévias do hospedeiro como a idade, sexo e comorbidades pré-existentes (ASAKURA; OGAWA, 2021)

Após a entrada do vírus no corpo do hospedeiro, ele é reconhecido pelo sistema imune inato por meio das células apresentadoras de antígeno (APC), como as células dendríticas e macrófagos, os quais compõem a linha de frente do sistema imunológico. As APCs possuem receptores de reconhecimento de padrões (PRR), que incluem os receptores Toll-like (TLRs), receptores NOD-like (NLRs), receptor RIG-I-like e outras moléculas que estão localizadas em várias partes das células hospedeiras, como membranas plasmáticas, lisossomas e citosol. Esses receptores reconhecem os padrões moleculares associados a patógenos (PAMPs), que são componentes estruturais dos vírus, como ácido nucleicos, proteínas, carboidratos e lipídeos. Posteriormente é induzida a sinalização, onde cada um dos PRRs pode ativar uma resposta biológica diferente (TURČIĆ *et al*., 2024).

As APCs apresentam antígenos do SARS-CoV-2 aos linfócitos T auxiliares (Th) CD4+ por meio do complexo principal de histocompatibilidade (MHC) de classe II, isso leva a secreção de IL-12, que aumenta o estímulo à ativação de células Th1 e em associação com o IFN-α, aumenta a expressão de MHC de classe I e a ativação de células *natural killer* (NK), aumentando a atividade da via de sinalização NF-kB, que induz a produção de mais citocinas pró-inflamatórias, como a IL-17. O papel destas citocinas é recrutar neutrófilos e monócitos para o sitio de infecção, ativar e induzir a

produção de várias outras citocinas e quimiocinas, como IL-1, IL-6, IL-8, IL-21 e TNF-β (MUSSABAY *et al*., 2024). Os interferons do tipo 1 (IFN-1) são conhecidos por atuarem em reações antivirais, desencadeando a liberação de enzimas que degradam o ácido nucleico viral e que produzem proteínas que bloqueiam a liberação de partículas virais por células infectadas. Contudo, o SARS-CoV-2 consegue neutralizar a resposta imune do IFN-1, sendo capaz de bloquear a síntese e degradar o mRNA que codifica o IFN-1 (CHUNG et al., 2024b).

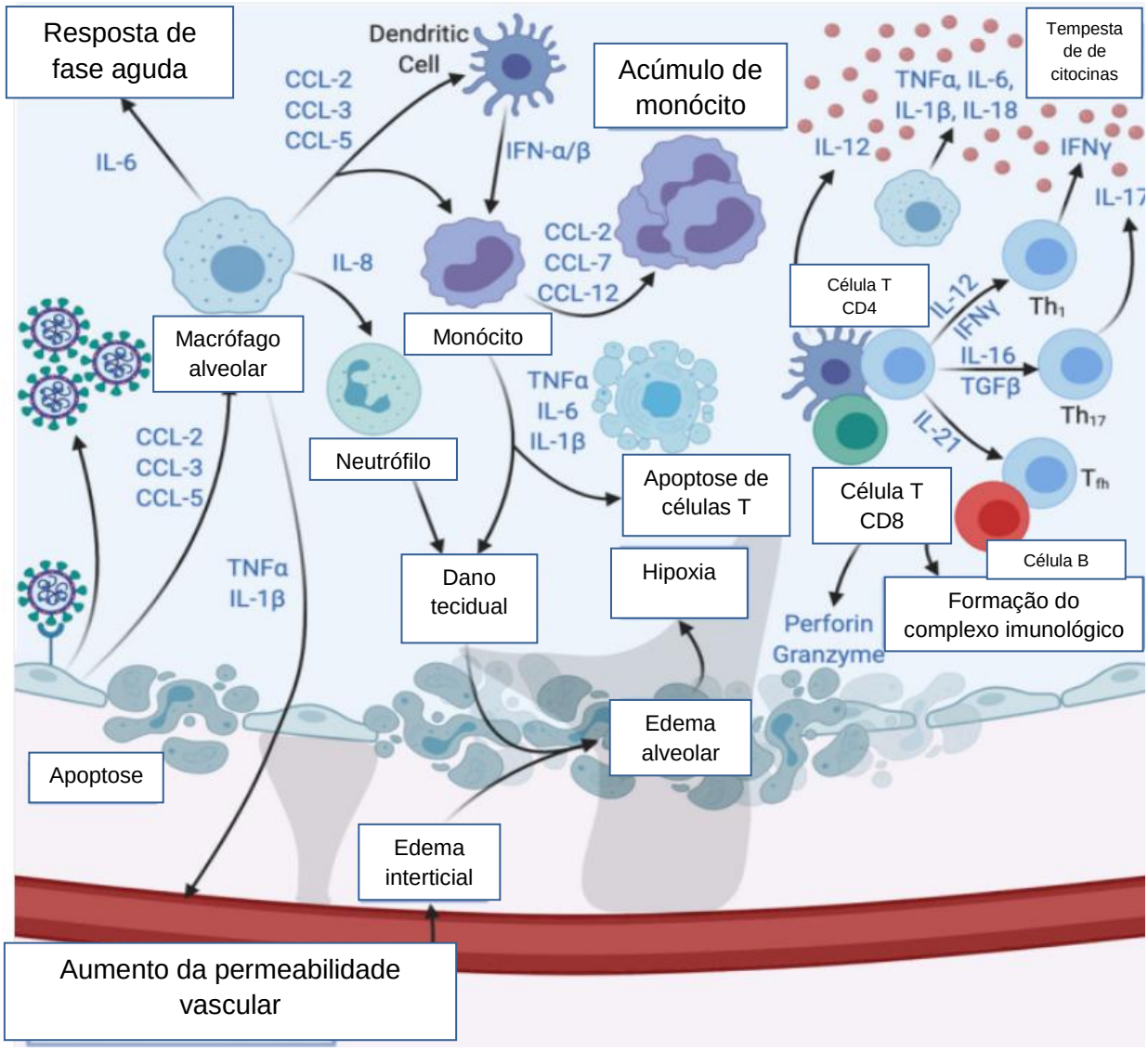
A principal responsável pelo agravamento dos casos de covid-19 é a estimulação excessiva de células efetoras na produção de citocinas como IFN-α, IFN-γ, IL-1β, IL-6, IL-12, IL-18, IL-33, TNF-α, TGF-β, e quimiocinas como CCL2, CCL3, CCL5, CXCL8, CXCL9 e CXCL-10, que causam uma hiperinflamação e conduzem a um quadro de síndrome da angústia respiratória aguda (SARA) (ATTIQ *et al*., 2024a).

A ação de células T CD8+ pode ser induzida pela ativação de células Th1 ou pela apresentação de antígenos pelas APCs, através do MHC de classe I. Ao serem ativadas, estas células sofrem expansão clonal, produzindo mais células T efetoras e de memória (FIGURA 5). Os linfócitos T CD8+ compõem um grupo específico de células T efetoras, que marcam e lisam células infectadas pelo SARS-CoV-2 por meio de múltiplos mecanismos, incluindo perforinas e granzimas (SUBBARAO MALIREDDI; SHARMA; KANNEGANTI, 2024).

A estimulação contínua dos vírus nas células TCD8+ causa uma diminuição da capacidade proliferativa e nas funções efetoras, induzindo a superexpressão de receptores inibitórios como o CD279, também conhecido como PD-1. O PD-1 é uma proteína presente na superfície das células linfoides e um membro da família CD28/CTLA-4, responsável pela regulação das células T, principalmente na modulação da apoptose (SONG *et al*., 2024).

O SARS-CoV-2 causa uma lesão tecidual, causando uma secreção exacerbada de citocinas pró-inflamatórias, levando ao recrutamento de granulócitos, macrófagos e outras células inflamatórias. O aumento da secreção de citocinas com o recrutamento de leucócitos causa uma resposta inflamatória sistêmica, chamada de tempestade de citocinas (ATTIQ *et al*., 2024b; CHENG *et a*l., 2023).

Figura 5 – Imunopatologia da covid-19 e a ação das células após infecção



Fonte: Cheng *et al.,* (2023). A entrada do vírus na célula causa um amplo recrutamento de moléculas do sistema imunológico causando uma cascata de reações para tenta combater o invasor ou até causar apoptose celular.

As células Th1 estimulam também a resposta imune humoral, através da produção de anticorpos específicos via células B dependentes de células T. A infecção por SARS-CoV-2 leva a produção de imunoglobulinas de cadeia M (IgM) específicas que podem ser observadas logo na primeira semana do início dos sintomas (fase aguda), durando até 12 semanas. Após uma média de 14 dias ocorrerá a soroconversão de anticorpos IgG, que permanecem detectáveis por períodos maiores Embora os períodos médios para produção de IgM e IgG já tenham sido estabelecidos, há frequentemente casos de covid-19 que não seguem um padrão,

com uma nova ocorrência de IgM após o surgimento de IgG, junto com IgG ou mesmo ausência de detecção (BELIK *et al*., 2024)

2.7 MARCADORES HUMANOS ASSOCIADOS AO QUADRO DE SAÚDE

2.7.1 ACE-2 e a covid-19

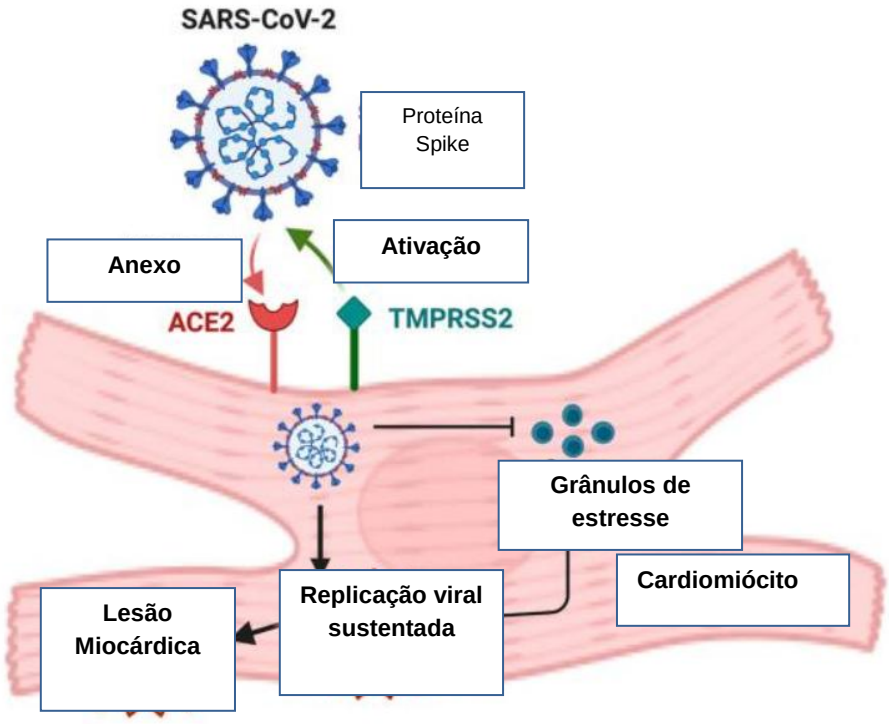
A ACE-2 está amplamente distribuída no corpo humano, ficando ligada à membrana de células dos tecidos de órgãos como o coração, rim e em menor extensão, nos pulmões, onde está concentrada em células alveolares do tipo II. A ACE-2, é uma monocarboxipeptidase responsável por degradar a angiotensina II, um peptídeo pró-hipertensivo e pró-fibrótico, para gerar a angiontensina 1-7, um produto vasodilatador e cardioprotetor, que desemprenha papel na regulação do Sistema Renina-Angiotensina-Aldosterona (SRAA; SANTOS *et al*., 2017). A partir da alteração na homeostase dos fluidos ou da pressão arterial, é iniciada a ativação do SRAA através da liberação da renina no sangue, causando a conversão do angiotensinogênio angiotensina I (Ang I), através da enzima conversora de angiotensina (ECA), em Ang II. Esse mecanismo é o responsável pela SRAA e assim causa a vasoconstrição e o aumento da pressão arterial (PANNUCCI *et al*., 2023).

O ACE-2 é o mais conhecido receptor do SARS-CoV-2. A proteína S do SARS-CoV-2 se liga ao ACE-2, permitindo a ligação com a célula alvo. Um estudo realizado por Kuba *et al., (*2005) mostrou que a inativação da ACE-2 reduz significativamente a infecção e replicação viral em camundongos infectados por SARS-CoV-2. Outro estudo avaliou o desenvolvimento da infecção causada por SARS-CoV-2 em camundongos e evidenciou que aqueles deficientes em TMPRSS2 foram protegidos do agravamento da doença (IWATA-YOSHIKAWA *et al*., 2019).

O gene *ACE-2* está presente no cromossomo X, e um dos polimorfismos mais investigados está associado a hipertensão e doenças cardíacas. Goméz *et al.,* (2020) relataram que o polimorfismo dessa região está associado ao agravamento, dependendo do estado da hipertensão do paciente. Isso ocorre devido a formação do complexo SARS-CoV-2/ACE-2 que permite que o vírus entre na célula hospedeira e produza uma regulação negativa na superfície da ECA2, causando uma perda de função cardioprotetora, e assim, uma perda no sistema cardiovascular. Esse

comprometimento causa uma resposta vasoconstritora com inflamação, hipertrofia e um estado pró-fibrótico, levando ao dano no miocárdio causado pela infecção viral (FIGURA 6) (PANNUCCI *et al*., 2023). Além dos efeitos causados no sistema cardiovascular, a perda das ações do ACE-2 no intestino causada pelo SARS-CoV-2 pode causar diarreia, náusea e vômitos, promovendo a inflamação intestinal e alteração do microbioma (PRINGLE; PHILP, 2023).

Figura 6 - Entrada do SARS-CoV-2 na célula com auxílio do TMPRSS2



Fonte: PANNUCCI *et al.,* (2023). A entrada do SARS-CoV-2 utiliza diretamente o ACE-2 e TMPRSS2. A utilização na maquinaria celular causa danos no tecido do hospedeiro e como o ACE-2 está amplamente distribuído, diversos órgãos podem ser afetados

2.7.2 TMPRSS2 e a covid-19

A TMPRSS2 é uma protease transmembrana do tipo II que é expressa em células epiteliais do trato respiratório, gastrointestinal, próstata e em outros órgãos. Foi identificado que o TMPRSS2 serve com uma via de entrada para o SARS-CoV-2, além de auxiliar na entrada de outros vírus. Sendo assim, inibir a atividade do

TMPRSS2 se torna essencial para inibir a ação viral (WETTSTEIN; KIRCHHOFF; MÜNCH, 2022).

O gene *TMPRSS2* humano possui 14 éxons e 13 íntrons e está presente no cromossomo 21 (21q22.3). O gene é conhecido na oncologia devido a sua fusão entre *TMPRSS2* e *ERG* (gene relacionado aos ETS – *ETS-related gene*), associadas ao câncer de próstata, onde em casos patológicos, o TMPRSS2 está expresso exageradamente e mais difuso (LUCAS *et al*., 2008).

Como citado anteriormente, o TMPRRS2 é bastante expresso nas vias aéreas dos seres humanos, e assim, apresenta um papel importante na entrada dos vírus respiratórios (LAPORTE; NAESENS, 2017). CHENG *et al.,* (2015) mostraram que o polimorfismo nas regiões rs2070788 foi significativamente associado à suscetibilidade à influenza A, onde o alelo rs2070788G conferiu um maior risco à forma grave da doença.

Para o SARS-CoV-2, devido a sua semelhança com os outros coronavírus, percebeu-se rapidamente que ele infecta rapidamente as células que expressam o ACE-2 e que ocorre devido a ativação da proteína S através do TMPRSS2 (SILVA *et al*., 2023a). Além do TMPRSS2, outros TMPRSS, como o TMPRSS11D, TMPRSS11E e TMPRSS11F, não ajudam na infecção pulmonar, mas podem contribuir para a disseminação de forma extrapulmonar do SARS-CoV-2 (HOFFMANN *et al*., 2021).

2.8 POLIMORFISMOS DE NUCLEOTÍDEO ÚNICO (SNP)

2.8.1 SNP *rs961360700* no gene ACE-2

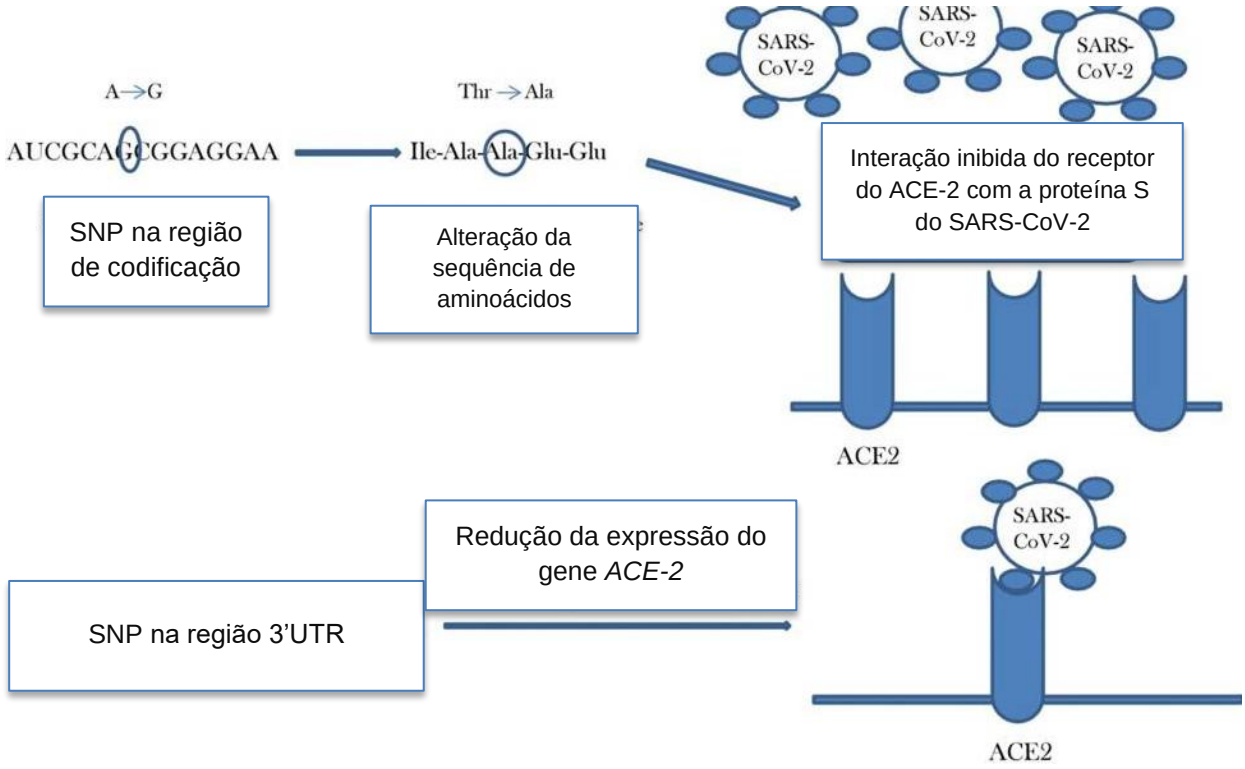
Diversos resíduos de aminoácidos são expressos no receptor ACE-2, apresentando uma relevância significativa na prevenção ou na promoção contra as infecções virais. Além disso, o ACE-2 humano é altamente polimórfico, apresentando 223 SNPs que podem resultar um mutações significativas, podendo influenciar na suscetibilidade ao SARS-CoV-2 (REN *et al*., 2022).

Diversos SNP’s de ACE-2 estão associados com a facilitação da entrada viral na célula do hospedeiro. A presença dos SNPs em regiões codificadoras pode causar uma alteração na sequência de aminoácidos diretamente no local de ligação do receptor à proteína do SARS-CoV-2 (FIGURA 7). Dessa forma, diversos SNPs nessa

região devem ser investigados de forma a até serem utilizados com alvos para possíveis novos fármacos (CHAUDHARY, 2020; TAHER *et al*., 2021).

Em um estudo realizado por Wooster *et al*. (2020), foram estudados dez polimorfismos no gene *ACE-2* onde cinco polimorfismos foram identificados como influenciadores da gravidade da doença, sendo eles: *rs4240157, rs6632680, rs4830965, rs1476524 e rs2048683*. Esses polimorfismos estavam diretamente relacionados com a maior expressão do ACE-2 no tecido dos pacientes, enquanto o *rs1548474* estava relacionado com baixa expressão, menor gravidade e menor hospitalização (WOOSTER *et al.*, 2020).

FIGURA 7 – Presença de SNP's de ACE-2 e alteração na suscetibilidade ao SARS-CoV-2



Fonte: Chaudhary (2020). A alteração na sequência de aminoácidos pode modificar a estrutura proteica envolvida na interação entre o SARS-CoV-2 e a célula do hospedeiro, podendo inibir a entrada viral.

O SNP *rs961360700* (C>T) não apresenta muitas caracterizações na literatura, mas já é descrito como um polimorfismo que pode impedir a interação direta entre o ACE-2 e a S1, devido à uma menor afinidade de ligação com a proteína *spike*, podendo reduzir significativamente a entrada do SARS-CoV-2 e assim a infecção (WANG *et al*., 2020).

2.8.2 SNP *rs2298659* no gene *TMPRSS2*

Diversos estudos investigaram a influência dos polimorfismos do gene *TMPRSS2* na suscetibilidade ou proteção ao SARS-CoV-2 (IRHAM *et al*., 2020; KOTSEV *et a*l., 2021; MONTICELLI *et al*., 2021; SENAPATI *et al*., 2020). O SNP *rs2070788* caracterizado pela substituição da guanina (G) pela adenina (A) foi investigado em países europeus, africanos e asiáticos, onde a presença do polimorfismo estava diretamente relacionada com a mortalidade elevada em pacientes europeus (ASSELTA *et al.*, 2020; SILVA *et al.*, 2022).

O SNP *rs2298659* (G>A, C) é o polimorfismo mais encontrado em na população mexicana e a análise “in sílico” avaliou que ele pode apresentar um impacto em nível do mRNA, podendo afetar o *splicing* e assim causar uma isoforma do TMPRSS2. A isoforma seria causada pela presença do alelo A do SNP, responsável pela criação de um sítio de ligação para SRp40, proteína de *splicing* que ajuda na formação da isoforma, que consequentemente ajuda a clivar a proteína S, facilitando a entrada do SARS-CoV-2 (POSADAS-SÁNCHES *et al.,* 2022).

3. JUSTIFICATIVA

A pandemia causada pelo SARS-CoV-2 representou um dos maiores desafios recentes para a saúde pública mundial. O Brasil figura entre os países mais afetados, sendo o sexto país em número de casos e o segundo em número acumulado de óbitos. (BRASIL, 2025). Apesar das medidas de contenção implementadas, a elevada vulnerabilidade social e do sistema de saúde brasileiro contribuiu para o agravamento do impacto da doença. Nesse cenário, sucessivas ondas de pandemia intensificaram a pressão sobre os serviços de saúde, especialmente sobre os profissionais que permaneceram na linha de frente do cuidado à população (SILVA *et al*., 2023b).

A infecção pelo SARS-CoV-2 depende de mecanismos moleculares bem estabelecidos: a proteína Spike do vírus interage com o receptor da enzima conversora de angiotensina 2 (ACE-2) e sofre clivagem pela protease transmembrana serina 2 (TMPRSS2), permitindo a entrada viral na célula hospedeira. Embora esses genes já tenham sua importância reconhecida na literatura, pouco se sabe sobre o papel específico de determinados polimorfismos de nucleotídeo único (SNPs) em sua expressão ou função. Em particular, os SNPs *rs961360700* (ACE-2) e *rs2298659* (TMPRSS2) já foram associados a outras condições clínicas e comorbidades, mas ainda não há evidências consolidadas sobre sua relação direta com a fisiopatologia da covid-19, com a evolução sintomática da infecção ou com o agravamento da doença (POSADAS-SÁNCHEZ et al., 2022; REN et al., 2021). Os genes escolhidos para esse trabalho já possuem sua importância para a entrada viral nas células do hospedeiro bem descritas na literatura. No entanto, os SNPs selecionados, apesar de estarem relacionados com outras doenças ou comorbidades, não é descrita a sua relação com a fisiopatologia na infecção por SARS-CoV-2, evolução da sintomatologia e/ou agravamento da doença.

Durante a primeira onda da pandemia, os profissionais de saúde constituíram um dos grupos mais vulneráveis devido à intensa exposição a elevadas cargas virais. Esse risco foi potencializado por condições adversas da época, como a escassez de equipamentos de proteção individual (EPIs), além da ausência de protocolos claros de descontaminação e contingência (BLACK *et al*., 2020; ZHANG *et al*., 2023). Diante desse contexto, é fundamental compreender como fatores genéticos, em especial sobre as variantes de genes diretamente envolvidos na entrada viral, que podem

influenciar a susceptibilidade, a evolução clínica e a gravidade dos casos entre esses profissionais.

Portanto, investigar o perfil genotípico e alélico dos SNPs selecionados em ACE-2 e TMPRSS2, correlacionando-os com dados clínicos, epidemiológicos e sociais dessa população, poderá fornecer subsídios relevantes para elucidar mecanismos associados à variabilidade clínica da covid-19. Além disso, os achados poderão contribuir para a compreensão de riscos individuais em futuras epidemias virais e para o direcionamento de estratégias de prevenção e manejo clínico em profissionais de saúde expostos a patógenos emergentes.

4 OBJETIVOS

4.1 OBJETIVO GERAL

Avaliar o perfil genotípico e alélico de SNPs dos genes *ACE-2* e *TMPRSS2* associados a sintomatologia, evolução clínica e a sua correlação com as caraterísticas clínicas e sociais de profissionais que foram expostos ao SARS-CoV-2 durante a primeira onda da pandemia da covid-19.

4.2 OBJETIVOS ESPECÍFICOS

- Verificar as frequências dos genotípicas e alélicas dos SNPs *rs961360700 ACE-2* e *rs2298659 TMPRSS2* na população do estudo;
- Caracterizar o perfil epidemiológico dos profissionais expostos ao SARS-CoV-2, de acordo com a sintomatologia apresentada e a gravidade dos casos de covid-19;
- Investigar a associação entre a evolução da covid-19 e os SNPs *rs961360700 ACE-2* e *rs2298659 TMPRSS2* nos indivíduos que adoeceram*;*
- Verificar a presença de comorbidades na população estudada e sua possível correlação com os SNPs *rs961360700 ACE-2* e r*s2298659 TMPRSS2*;
- Realizar análises *in sílico* para identificar potenciais alterações funcionais decorrentes dos SNPs rs*961360700 ACE-2* e *rs2298659;*
- Elaborar e publicar uma revisão sistemática sobre a influencia dos polimorfismos no gene *TMPRSS2* na susceptibilidade ao SARS-CoV-2.

5. MATERIAL E MÉTODOS

5.1 DELINEAMENTO DO ESTUDO

Trata-se de um estudo observacional analítico do tipo coorte retrospectiva. A população do estudo foi composta por profissionais de diversas áreas, incluindo serviços gerais, administração e equipe de saúde, que atuaram em unidades assistenciais que atenderam pacientes com covid-19 durante a primeira onda da pandemia.

As instituições de saúde incluíram: Policlínica Metropolitana de Belém, Hospital Jean Bittar; Hospital das Clínicas Gaspar Vianna; Hospital Universitário João de Barros Barreto; Hospital Adventista de Belém; Hospital Dom Vicente Zico; Hospital de Saúde da Mulher; Instituto de Hematologia e Hemoterapia de Belém; Centro Psicossocial e Secretaria de Saúde Pública do estado do Pará.

Os participantes foram expostos ao SARS-CoV-2 no período compreendido entre 21/04/2020 a 30/06/2020, período definido como a primeira onda pandêmica no contexto local. A inclusão na coorte baseou-se na atuação ativa desses profissionais nas unidades que prestaram atendimento a pacientes confirmados ou suspeitos de covid-19 nesse intervalo.

5.2 CRITÉRIOS DE INCLUSÃO E EXCLUSÃO

Foram incluídos os profissionais que tiveram contato com pessoas sintomáticas infectadas pelo SARS-CoV-2, independentemente do número de horas ou dias trabalhados nas unidades em atendimento a covid-19 entre 21/04/2020 e 30/06/2020 e que aceitaram participar do estudo, onde assinaram um Termo de Consentimento Livre Esclarecido (TCLE) (APÊNDICE 4). Para serem considerados nos critérios de contato com sintomáticos, as manifestações sintomáticas dos pacientes deveriam ter ocorrido na primeira onda da pandemia (BOAVENTURA et al., 2023).

Foram excluídos os indivíduos que apresentavam doenças autoimunes que poderiam influenciar no aparecimento de sintomas.

5.3 CARACTERIZAÇÃO DA AMOSTRA

A inclusão dos participantes foi por conveniência, sendo os profissionais convidados a participar do estudo. Portanto, foram incluídos no estudo 213 profissionais que tiveram contato com pessoas sintomáticas infectadas pelo SARS-CoV-2

Para a caracterização das amostras, os participantes foram divididos em dois grandes grupos: indivíduos assintomáticos (AS) e sintomáticos (SI).

O grupo de indivíduos assintomáticos (AS) foi constituído por indivíduos que apesar do contato com pacientes infectados, não apresentaram sintomas característicos de COVID-19 no período de março de 2020 a julho de 2020.

Enquanto que o grupo dos indivíduos sintomáticos (SI), compreendeu os participantes que relataram a presença de pelo menos dois sintomas característicos de covid-19.

Os pacientes do grupo SI foram divididos em: indivíduos sintomáticos sem comprometimento pulmonar (SSP), com a presença de dois sintomas, mas com a ausência de lesões pulmonares e pacientes sintomáticos com comprometimento pulmonar, sendo composto por pacientes com pelos menos dois sintomas e com a presença de comprometimento pulmonar de no mínimo 10%, podendo ser analisado pela tomografia do tórax.

5.4 COLETA DOS DADOS

Aos participantes foram aplicados questionários epidemiológicos estruturados, no qual foram coletadas informações sobre a sintomas (incluindo tempo de sintomas), severidade, internação, comorbidades, características sociodemográficas e comportamentais (APÊNDICE 5).

5.5 ANÁLISE DAS TOMOGRAFIAS DO TÓRAX

A análise dos exames de imagem foi executada por uma médica pesquisadora do grupo de pesquisa, sendo ela a responsável pela interpretação das tomografias e

diagnostico de possíveis danos causados pelas infecções respiratórias. As imagens foram cedidas pelos profissionais que realizaram durante o período da primeira onda. A partir disso, os achados foram classificados de acordo com o grau de acometimento do parênquima pulmonar, sendo avaliadas as “opacidades em vidro fosco” de acordo com a sua distribuição, sendo classificado em multifocal, bilateral, posterior, periférica e basal.

Para a padronização dos laudos com os achados nas tomografias, foi utilizado o CO-RADS, sistema específico para covid-19 disponibilizado pela Radiological Society of the Netherlands (SIMPSON *et al*, 2020; PENHA *et al*., 2021)

5.6 COLETA DAS AMOSTRAS CLÍNICAS

As coletas das amostras foram realizadas entre os meses de janeiro de 2021 a janeiro de 2023, nos centros de saúde onde os participantes trabalhavam. Foram coletadas amostras de sangue total dos profissionais por punção venosa em tubos para coleta de sangue VACUETTE® (Greiner Bio-One) de 5 mL, contendo EDTA. Após a coleta, o material era armazenado em um isopor com gelo reutilizável em gel e foram transportadas para serem estocadas à -20 ºC na Seção de Bacteriologia e Micologia (SABMI) no Instituto Evandro Chagas (IEC).

5.7 PROCEDIMENTOS LABORATORIAIS

5.7.1 Extração de DNA

O DNA foi extraído através do kit QIAmp ® DNA Blood Mini kit (QIAGEN) seguindo o protocolo com as especificações do fabricante (ANEXO 1).

5.7.2 Tipificação dos SNPs dos genes ACE-2 e TMPRSS2

Para a amplificação dos fragmentos de DNA das regiões genômicas com polimorfismo, foi realizada a técnica de PCR com a utilização de iniciadores específicos que foram desenhados no programa Prime3Plus versão 3.3.0

(https://www.primer3plus.com/index.html) a partir das regiões genômicas que estão depositadas no GenBank (NCBI) (Quadro 1).

Quadro 1- SNPs e iniciadores utilizados no estudo para a análise dos polimorfismos

Gene	SNP	Sequência dos Primers
ACE 2	*rs961360700*	F: 5’ GGGTGCAACCATCCCCATTA 3’ R: 5’ GCCTGGGATGCACAGAGAAT 3’
TMPRSS2	*rs2298659*	F: 5’GCATGAGCGCACTTGATGTC 3’ R: 5’GTCCCTGAGTGGTGTCAGTC 3’

Fonte: Própria autora (2025). F: Forward; R: Reverse

A amplificação foi realizada com o uso de Platinum Taq em um termociclador utilizando as seguintes configurações: desnaturação inicial a 95 ºC por 1 minuto, seguidos de 35 ciclos de desnaturação de 95 ºC a 30 segundos, anelamento a 65 ºC por 30 segundos e extensão por 72 ºC a 1 minuto, finalizando com uma extensão final à 72 ºC por 10 minutos.

Os produtos foram analisados através da visualização em gel de agarose UltraPure™ Agarose (Invitrogen- Thermo Fisher Scientific) a 2% contendo o corante de DNA SYBR® Safe Stain (Edvotek), sob as condições de 120 v, 120 mA, 100 W com duração de 30 minutos, sendo visualizados em transiluminador.

Posteriormente, os produtos de PCR obtidos foram submetidos a um processo de purificação utilizando o kit *EasyPure*® Genomic DNA Kit (TransGen Biotech), seguindo o protocolo do fabricante.

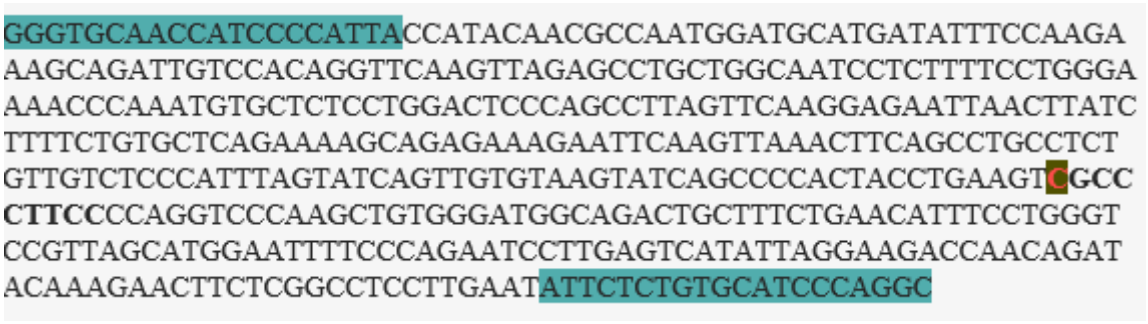
5.7.3 Sequenciamento de Sanger

O sequenciamento pelo o método Sanger foi realizado utilizando o kit BigDye ™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems®), de acordo com o protocolo do fabricante. Posteriormente, foi utilizado o kit BigDye XTerminator™ Purification Kit (Applied Biosystems®), seguindo as orientações do fabricante. Após esta etapa, foram montadas as placas e assim as reações foram colocadas no sequenciador ABI 3130 Genetic analyzer (Applied Biosystems®).

5.7.4 Análise pós sequenciamento e detecção de SNP

Os fragmentos obtidos foram comparados com as sequencias modelo obtidas no Basic Local Alignment Search Tool (BLAST), disponível no *National Center for Biotechnology Information* - NCBI (https://www.ncbi.nlm.nih.gov/). Posteriormente, baseado na localização do primer, foi identificada a posição do SNP e assim anotada qual nucleotídeo foi detectado na amostra analisada (FIGURA 8; FIGURA 9).

Figura 8 – Localização da região do SNP do gene *ACE-2 rs961360700*



Fonte: Própria autora (2025). Em realce azul é mostrada a localização dos iniciadores *forward* e *reverse*; em vermelho é mostrado o nucleotídeo polimórfico.

Figura 9 – Localização da região do SNP do gene *TMPRSS2 rs2298659*

Fonte: Própria autora (2025). Em realce vermelho é mostrada a localização dos iniciadores *forward* e *reverse*; em verde é mostrado o nucleotídeo polimórfico.

5.8 ANÁLISE ESTATISTÍCA

Os dados extraídos foram organizados em planilha do Microsoft Office Excel. A avaliação do poder amostral post-hoc da amostra N foi realizada utilizando o software G*Power versão 3.1.9.7 com um teste qui-quadrado de ajuste, com

probabilidade de erro α de 0,05 e tamanho do efeito de 0,3, considerando a aceitabilidade do poder amostral de 0,8.

A normalidade das variáveis foi avaliada pelo teste de Kolmogorov-Smirnov. Para a análise correlacional entre as variáveis dependentes e independentes, foi utilizada a regressão logística multinomial, comparando cada categoria com a categoria de referência (Grupo AS) para os parâmetros de sintomatologia e gravidade da COVID-19. A probabilidade (p) ≤ 0,05 foi considerada estatisticamente significativa.

Em relação aos dados genotipados da coorte de Belém, as frequências genotípicas foram testadas para o Equilíbrio de Hardy-Weinberg (EHW) utilizando o teste qui-quadrado ($\chi 2$) com $p < 0,001$ como ponto de corte para o nível de significância por meio do SNPs Analyzer v. 1.1.82. As frequências dos genótipos e alelos dos SNPs foram analisadas em conjunto com os dados epidemiológicos, também de forma multinominal, para verificar e mensurar os efeitos.

5.9 CONSIDERAÇÕES ÉTICAS

O presente projeto faz parte de um projeto de pesquisa do tipo “guarda-chuva” intitulado “Análise da resposta ao SARS-CoV-2 em relação aos achados radiológicos e/ou à susceptibilidade genética individual” que foi submetido ao comitê de ética e pesquisa da Universidade do Estado do Pará – Centro de Ciências Biológicas e da Saúde em agosto de 2020, sendo aprovado com o número do CAAE: 38113620.1.0000.5174 (ANEXO 2).

6. RESULTADOS

ARTIGO 1 – Artigo publicado no International Jourrnal of Genetics.

International Journal of Genetics
ISSN: 0975-2862 & E-ISSN: 0975-9158, Volume 14, Issue 4, 2022, pp.-850-855.
Available online at https://bioinfopublication.org/pages/jouarchive.php?id=BPJ0000226

Review Article

THE ASSOCIATION BETWEEN *TMPRSS2* GENETIC POLYMORPHISMS AND THE SUSCEPTIBILITY AND SEVERITY OF COVID-19: A SYSTEMATIC REVIEW

SILVA C.S.[1,3*], SILVA M.J.A.[2,3], LIMA K.V.B.[1,2,3], FROTA C.C.[4], SARDINHA D.M.[1,3], LIMA L.N.G.C.[1,2,3]

[1]Postgraduate Program in Parasitic Biology in the Amazon (PPGBPA), University of State of Pará (UEPA), Brazil
[2]Program in Epidemiology and Health Surveillance (PPGEVS) of the Evandro Chagas Institute (IEC), Ananindeua, Pará, Brazil
[3]Bacteriology and Mycology Section of the Evandro Chagas Institute (IEC), Ananindeua, Pará, Brazil
[4]Department of Pathology and Legal Medicine, Faculty of Medicine, Federal University of Ceará, Fortaleza, CE, Brazil
*Corresponding Author: Email - karolinysoares2303@gmail.com

Received: October 01, 2022; Revised: October 26, 2022; Accepted: October 27, 2022; Published: October 30, 2022

Abstract: COVID-19, caused by the SARS-CoV-2 virus, is a highly transmissible disease that has a variety of symptoms. The presence of genetic polymorphisms in genes as ACE-2 and *TMPRSS2* is directly associated with the susceptibility and severity of COVID-19. The objective of this work is to analyze which polymorphisms in the *TMPRSS2* gene are associated with the progression of COVID-19. We identified 35 SNPs associated with disease progression. The high expression of *TMPRSS2* in the lungs was associated with the presence of polymorphisms such as rs383510, rs469390 and rs464397. rs12329760 was the polymorphism most studied, where the presence of the T allele was related to protection against COVID-19. The SNPs found play an important role in determining the prognosis of the disease.

Keywords: *COVID-19, SARS-CoV-2, Polymorphism, TMPRSS2, Expression*

Citation: Silva C.S., *et al.*, (2022) The Association Between *TMPRSS2* Genetic Polymorphisms and The Susceptibility and Severity of Covid-19: A Systematic Review. International Journal of Genetics, ISSN: 0975-2862 & E-ISSN: 0975-9158, Volume 14, Issue 4, pp.- 850-855.

Academic Editor / Reviewer: I S Chakrapani, Prof Dr Suleyman Cylek

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the pathogen responsible for causing coronavirus disease 2019 (COVID-19). This virus was discovered in 2019 in Wuhan, China, and showed high infectivity, spreading rapidly throughout the world, causing a global pandemic [1].

SARS-CoV-2 spreads primarily through respiratory droplets when people talk, cough, sneeze, or through direct contact with the virus on contaminated surfaces [2]. The pathogen mainly infects the respiratory system, where infected individuals may present several clinical symptoms, such as coughing, fever, dyspnea, skin manifestations, and muscle pain [3]. Approximately 15% of people with COVID-19 develop the severe form of the disease, which can progress to acute respiratory distress syndrome (ARDS), severe pneumonia, kidney damage, and death [4]. The severity of COVID-19 appears to be affected by several risk factors, such as advanced age, cardiovascular disease, obesity, and diabetes [5].

Coronaviruses (CoVs) belong to the order Nidovirales, family Coronoviridae, and subfamily Orthocoronavinae, being subdivided into four genera: Alphacoronavirus, Betacoronavirus, Gammacoronavirus and Deltacoronavirus [6]. SARS-CoV-2 belongs to the group of β -coronaviruses with enveloped, single-stranded, positive-sense RNA [7]. The complete genome of SARS-CoV-2 is approximately 29.9 kb, with a GC content of 38% and 12 open reading frames (ORFs) [8]. CoVs have a spherical outline, with a diameter of the virion ranging from 60 to 140 nm and have characteristic structural proteins: spike (S), membrane (M), envelope (E) and nucleocapsid (N) [9]. In addition, they have 9 to 12 nm spicules that project onto the viral surface, having the appearance of a crown [10].

The entry of SARS-CoV-2 into cells occurs through the angiotensin converting enzyme 2 (ACE2), which is expressed in several organs, such as the lungs, intestine, heart, and kidneys [11]. Binding occurs by the S-protein binding domain of the virus that attaches tightly to human ACE2. Target cell proteases activate the S protein, causing cleavage in the S1 and S2 subunits, allowing fusion between viral and cellular membranes [12].

SARS-CoV-2 uses transmembrane serine protease 2 (*TMPRSS2*) as the enzyme responsible for cleaving and activating the S protein in the fusion process of the human membrane [13].

TMPRSS2 is a well-studied protein because it is involved in the development of prostate cancer due to the overexpression of specific transcription factors transforming erythroblasts (ETS), such as ERG [14]. The *TMPRSS2* gene, responsible for encoding *TMPRSS2*, is related to pathological and physiological processes, such as inflammatory responses, invasion of tumor cell, and fertility. and pain [15]. The *TMPRSS2* gene is found on chromosome 21q22.3, has 15 exons, and an open reading frame of 492 amino acids. This gene harbors androgen-responsive elements in its 5' UTR [16]. Testosterone and dihydrotestosterone are responsible for gene regulation through stimulation of androgen receptors [17].

TMPRSS2 has been shown to be crucial for the entry of SARS-CoV-2 into cells. From this, several studies sought to investigate the hypothesis that genetic variability and expression of the *TMPRSS2* gene are associated with susceptibility to COVID-19 [18].

Host genetic polymorphisms are known to play an important role in determining susceptibility or resistance to viral infections [19]. The host genes play an important role in the entry and replication of SARS-CoV-2 in cells and in the immune response, where the combination of genes may be associated with the pathogenesis of COVID-19 [20]. Studies have shown that gene polymorphisms involved in viral entry into human cells, such as ACE2 and *TMPRSS2*, are associated with susceptibility and severity to disease [21, 22].

Knowing the importance of polymorphisms in genes involved in the entry of SARS-CoV-2 into cells and in determining the susceptibility and severity of COVID-19, the following question arises from the research described here: 'Which SNPs in the *TMPRSS2* gene are associated with susceptibility, aggravation, or protection against COVID-19?'

Materials and Methods

Study design

This is a systematic review of the literature that aims to collect evidence on the reported correlations between the *TMPRSS2* gene polymorphisms and the progression of COVID-19. The review was conducted in accordance with the Preferred Reporting items for Systematic Reviews and Meta-Analyses Protocols (PRISMA Protocol) guidelines. The study followed the following training steps: 1) Elaboration of the guiding question; 2) Stipulation of inclusion and exclusion criteria; 3) Choice of articles; 4) Analysis of articles; 5) Interpretation, discussion, and presentation of the review. The PICO strategy (Population, Intervention, Comparison, Outcome) was used, following the following characteristics: Population: patients with COVID-19; Intervention: Assess the SNPs of the *TMPRSS2* gene in COVID-19 patients; Comparison: COVID-19 and the described SNPs of the *TMPRSS2* gene; Outcome: protection, susceptibility, or severity of COVID-19; Outcome: progression of COVID-19. From this, the following question was generated: What existing SNPs in the *TMPRSS2* gene are associated with susceptibility/worsening or protection against COVID-19?

Data sources

The terms used in the search based on Medical Subject Headings (MeSH) were: "COVID-19"; "SARS-CoV-2"; "*TMPRSS2*" and "Polymorphism". The articles were searched using the Boolean operator 'AND', in the following databases: National Library of Medicine National Institutes of Health of the USA (PubMed), Latin American and Caribbean Literature in Health Sciences (LILACS), Web of Science, and Scientific Electronic Library Online (SCIELO).

Study eligibility criteria

The selected studies were published from the beginning of the publications until June 2022, available complete, clinical studies, comparative studies, cross-sectional studies, case-control studies, cohort studies (prospective and retrospective), in vivo, in vitro trials, and review article narratives. Articles that identified SNPs of the *TMPRSS2* gene associated with susceptibility or progression of cases of COVID-19 were selected. As exclusion criteria, systematic review articles, abstracts, letters to the editor, and those that had subjects not relevant to the research were not included.

Data collection and extraction

Data collection was conducted in July 2022. Two researchers read the articles and independently extracted data from the publications following a predefined protocol. Information such as title, used method (case-control or cohort study), database, and relevant results were collected. The collected data were transferred to a Microsoft Office Excel spreadsheet.

Results

The search strategy resulted in 137 articles, and 81 were excluded since they were duplicates or were not relevant to the investigation. The abstracts of each of the remaining articles were evaluated and 17 were also excluded. At the end, full articles were evaluated and 21 met the inclusion criteria and therefore were added to the search [Fig-1]. The results showing the included studies are summarized in [Table-1]. Thus, 35 SNPs associated with susceptibility, symptom development, and severity of patients infected with SARS-CoV-2 were identified.

Discussion

COVID-19 was initially a disease associated with respiratory symptoms. Subsequently, it was found that the symptoms and severity of the cases can vary, where some patients may have mild symptoms (cough, headache, and fever) or develop acute respiratory illness, which can cause the death of these individuals [44]. The mechanisms involved in the susceptibility and evolution of COVID-19 cases are still unclear; however, the influence of host genetic polymorphisms on the immune response and its association with the severity of the disease are already known [45]. Studies have shown that regulatory genetic variants influence the expression of the *TMPRSS2* gene and change susceptibility to COVID-19, as well as worsening cases [24,45,47].

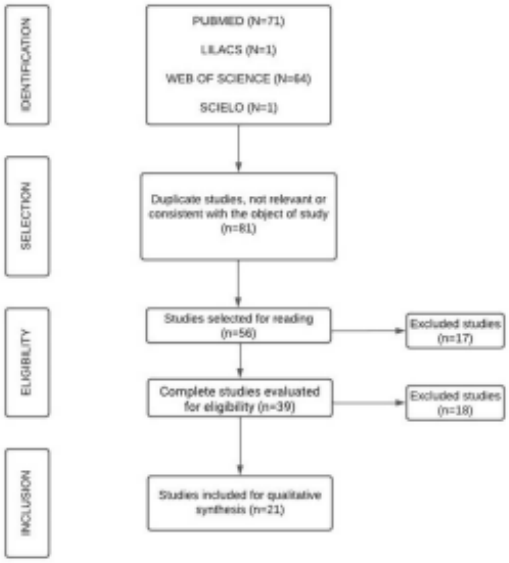


Fig-1 Flowchart of procedures for identification, selection, eligibility, and inclusion of studies for analysis. Belém, PA, Brazil (2022).

This review identified several *TMPRSS2* polymorphisms that are associated with the progression of COVID-19 cases. From the analysis of the selected articles, 35 SNPs were identified that influence the susceptibility, protection, and aggravation of COVID-19 cases. The rs12329760 SNP is described as a substitution of cytosine (C) for thymine (T). This polymorphism was analyzed in several studies, and it was found in most studies that the presence of the mutant T allele is considered protective against the infection and progression of COVID-19, while the wild C allele was associated with a greater number of hospitalizations of patients with COVID-19 [24, 26, 30, 31, 33, 34, 35, 37, 39, 42]. Computational analysis of changes caused by rs12329760 (C>T) in *TMPRSS2* structure shows a change from valine to methionine at position 197 (V197M), helping to form a pocket protein in the structure of *TMPRSS2*, which disfavors its binding to the SARS-CoV-2 S protein [35]. The importance of this SNP is evidenced in the analysis of its frequency in different populations and the advance of COVID-19 in these places. In a study conducted on Italian patients by Monticelli *et al.* (2021), a low frequency of the rs12329760 allele T was detected in this population, where it was suggested that the low frequency of this allele in the SNP had an influence on the high impact of the first wave of the pandemic in Italy [31].

Asselta *et al.* (2020) also analyzed the association of genetic polymorphisms and the severity of COVID-19 among Italians. The SNP rs12329760 and two distinct haplotypes were identified, showing differences in mutation frequency between Italians and East Asians. The presence of rare alleles in both haplotypes was associated with the induction of higher levels of *TMPRSS2*. The first haplotype (SNPs rs463727, rs34624090, rs55964536, rs734056, rs4290734, rs34783969, rs11702475, rs35899679, and rs395041537) was seen to be androgen regulated, which could help explain the severity of COVID-1 virus cases in the Italian population. SNPs belonging to the second haplotype (rs2070788, rs9974589, rs7364083) were correlated with a higher susceptibility to viral infections such as influenza A (H1N1) [24].

Different studies have investigated the presence of other SNPs in populations and analyzed whether the presence of these variants was involved in the susceptibility and worsening of COVID-19 cases [26, 27, 31, 32, 38, 41]. The SNP rs2070788 is characterized by the substitution of a guanine (G) for an adenine (A) and has been investigated in African, American, European, and Asian populations. The frequency of the G allele was highly detected in Americans and Europeans, compared to the frequency of Asians and Africans, and may be associated with high mortality among American and European populations.

Table-1 *Characteristics of the studies included in the systematic review*

Kind of study	SNP	Results	Reference
Case control	rs75603675; rs61735794; rs61735792	The SNP rs75603675 (G>T) was the most frequent variant among those infected with SARS-CoV-2. The SNP rs61735794 (A allele) and rs61735792 (T allele) were also significantly detected in patients with COVID-19.	[23]
Descriptive study	rs12329760 (p.Val160Met) ; Haplotype 1: rs463727, rs34624090, rs55964536, rs734056, rs4290734, rs34783969, rs11702475, rs35899679 and rs35041537; Haplotype 2: rs2070788, rs9974589, rs7364083	SNP rs12329760 (p.Val160Met) was detected at a higher frequency in East Asians than in Europeans. The rare alleles of these haplotypes found, associated with the induction of higher levels of TMPRSS2, are more frequent in the Italian population than in the East Asian population.	[24]
Descriptive study	SNP rs2298659	SNP rs 2298659 (C allele) showed a strong correlation with the fatality rate of COVID-19.	[25]
Cohort	rs7560367; rs200291871, rs12329270	A lower frequency of the variant allele of SNPs rs7560367 (T allele), rs12329270 (A allele) and rs200291871 (C allele) was observed in Italian COVID patients compared to the allele frequency of these SNPs in the European population.	[26]
Descriptive study	SNP rs469390, rs2070788, rs383510 and rs464397	The four SNPs were associated with differential expression of TMPRSS2.The A allele and the AA genotype of SNP rs469390 showed higher gene expression in the lung. The T alleles of rs464397 and rs383510, associated with increased expression of the *TMPRSS2* gene in the lungs, were exhibited with a lower frequency in East Asian populations. The frequency of the G allele for rs2070788 was also lower in the East Asian population compared to the European, African and American population.	[27]
Revision	rs2070788, rs383510	The GG genotype of rs2070788 increases the expression of *TMPRSS2* in the lungs. Regarding to rs383510, the T allele is associated with increased expression of *TMPRSS2* in the lungs.	[28]
Descriptive study	rs12329760	The evaluation of the structural stability of the protein predicted that rs12329760 (T allele) is responsible for destabilizing TMPRSS2, thus compromising the binding affinity of TMPRSS2 to the spike protein of SARS-COV-2 and ACE-2.	[29]
Cross-sectional study	rs12329760	Patients with the SNP rs12329760 CC genotype of the had an association with symptomatic or severe COVID-19; the C allele was associated with a higher viral load and a higher number of deaths.	[30]
Descriptive study	rs12329760	SNP rs12329760 was detected at lower values for East Asians, Finns and Africans. The association of the low-frequency T allele of rs12329760 in the population with the high impact of the first wave of the epidemic in Italy has been suggested.	[31]
Revision	rs713400, rs112657409, rs119110678, rs77675406	The frequency of the T allele at rs713400 influences the expression of TMPRSS2 and is considered higher among Asians than among Europeans, Africans and South Asians.	[32]
Descriptive study	rs781089181, rs570454392, rs867186402, rs12329760, rs118518290, rs762108701	The six variants analyzed were identified decreasing the stability of *TMPRSS2*. The SNPs rs781089181, rs570454392, and rs867186402 were in a highly conserved region and were responsible for crucial changes in the protein. Analyzing the native structure of the protein, 5 disordered regions were formed due to the rs12329760 and rs118518290 variants.	[33]
Descriptive study	rs12329760	The variant rs12329760 (T allele) was predicted to be deleterious and harmful to the expression of *TMPRSS2*, decreasing the stability of the protein.	[34]
***In silico* / descriptive study**	rs12329760 and rs75603675	The SNP rs12329760 (T allele) was considered deleterious by 3 tools used. A new major pocket protein caused by rs12329760 (T allele) was predicted, affecting the structure of TMPRSS2, and affecting its role in the entry of SARS-COV-2. SNP rs75603675 (A allele) has demonstrated the ability to increase protein disorder and influence TMPRSS2 function in facilitating SARS-COV-2 entry.	[35]
Case-control	rs383510	Carriers of the CC genotype in the rs383510 SNP had a 1.73-fold increased risk of infection for SARS-CoV-2.	[36]
Revision	rs12329760	It was observed that the number of cases and high mortality is associated with a low prevalence of the rs12329760 SNP (T allele), this SNP being significantly associated with genetic inactivation of TMPRSS2. From this, the presence of the T allele in rs12329760 is considered a protective factor against infection and disease progression.	[37]
Descriptive study	rs2070788	Significant positive association between the G allele of SNP rs2070788 and the mortality rate of COVID-19 among populations with a high frequency of this variant, for example, the Indian population. PNS eQTL analysis shows that the GG genotype tends to influence significantly higher expression of the *TMPRSS2* gene in the lung, increasing vulnerability to COVID-19. This variant was found at a high frequency in Americans and at a lower frequency in Africans and East Asians. This may be associated with high mortality among the American population and low severity among Asians.	[38]
Case-control	rs12329760, rs17854725,rs75603675	The T allele of rs12329760 conferred an increased risk of the individual being infected with SARS-CoV-2. The TT genotype was associated with a higher incidence of the severe form of the disease and the TC genotype was associated with an increase in the lesions seen on computed tomography. The presence of the G allele in SNP rs17854725 (A>G) was associated with greater susceptibility, while deaths were more frequent in carriers of the AG genotype. The AA genotype of rs75603675 appears to reduce the risk of a severe form of COVID-19 compared to the CC genotype. The GG genotype of rs4303795 increases the severe form of the disease and the occurrence of lesions.	[39]
Case-control	rs75603675	The presence of snp rs756039675 (T allele) was significantly associated with increased susceptibility to COVID-19 or increased risk of severe disease.	[40]
Descriptive study/ and systematic review	rs2070788, rs383510	The frequency of the G risk allele of the *TMPRSS2* rs2070788 gene was lowest among Africans and highest among Americans. The SNP rs383510 (T allele), responsible for also increasing the expression of *TMPRSS2*, was more frequent in the European population, while the African population was also the least frequent.	[41]
Descriptive study	r s3787946, rs9983330, rs12329760, rs2298661 and rs9985159	The SNPs rs3787946 (C allele), rs9983330 (G allele), rs12329760 (T allele), rs2298661 (A allele) and rs9985159 (T allele) were less recurrent in hospitalized patients, suggesting their protective role. Furthermore, 4 SNPs associated with the reduction of TMPRSS2 in lung tissues were identified: rs9983330 (G allele), rs12329760 (T allele), rs2298661 (A allele) and rs9985159 (T allele).	[42]
Cross-sectional study	rs75603675, rs12329760	The variant rs12329760 had no effect on the severity of the case. Analysis of SNP rs75603675 showed that the CC or CA genotypes compared to AA were associated with more severe COVID-19.	[43]

This is due to the influence of the G allele and the GG genotype on increased expression of *TMPRSS2* in the lungs [27, 38, 41]. The intervention of this genotype was confirmed through the analysis of quantitative expression traits loci (eQTLs) that reported the increase in the expression of *TMPRSS2* in the lungs and its influence on the increase in the expression of MX1, a gene involved in the viral response [28, 48]. Cheng *et al.* (2015) reported that this genotype was associated with a 2x greater risk of developing Influenza A in severe form, due to its ability to increase the expression of TPMRSS2 in the lungs, facilitating viral entry and thus increasing the susceptibility of the individual [49].

SNPs rs713400, rs383510, and rs753675 were also analyzed in the same ethnic groups mentioned above. The SNP rs713400 (C>T) was investigated in a single study, which showed that the presence of the T allele directly interfered with *TMPRSS2* expression. This mutation was identified with high frequency in the East Asian population, a region where the overall case fatality rate (CFR) is significantly lower (3.6%) compared to the global CFR (6.9%), which may suggest a protective character of that allele over the population [32]. In this same study, it was identified that the SNPs rs112657409, rs119110678 and rs77675406 also had a significant influence on the expression of *TMPRSS2*. The authors did not show which variants were involved in the infection, so further studies are needed for full clarification.

The SNP rs383510 (T>C) is located within a functional region with potentiating activity in the expression of *TMPRSS2*, where the presence of the T allele exhibited a higher transcriptional level than the C allele [47]. Database analysis of eQLTs showed that the TT genotype increases lung tissue gene expression [28]. The T allele was detected less frequently in the Asian population and was frequently identified in residents of European countries that had a high mortality rate of approximately 15% during the first wave of the pandemic [27, 41]. The study conducted by Ihram *et al.* (2020) analyzed allele frequencies in European, African, American, East Asian, and Southeast Asian populations. Again, the T allele of rs383510 was associated with increased expression of the *TMPRSS2* gene. However, contrary to the others, Schönfelder *et al.* (2021) found a significant 1.73-fold increase in the risk of SARS-CoV-2 infection in carriers of the CC genotype [36]. SNP rs469390 (G>A) was identified as a missense mutation and is located within an exonic region. The study by Irham *et al.* (2020) detected the influence of the A allele and the AA genotype of the rs469390 SNP with the highest gene expression in the lung. The presence of the heterozygous AG genotype was associated with intermediate expression and the GG genotype with a lower expression, making it evident that the presence of the mutant A allele makes the individual more susceptible to COVID-19. In this study, all SNP alleles related to the highest expression of *TMPRSS2* in the lungs, including the wild-type T allele of rs464397 (T>C), were detected with low frequency in the East Asian population, indicating a possible protection for these individuals [27].

In a study conducted in Spain, rs75603675 (G>T/ p.Gly8Val) was the most frequent SNP of the *TMPRSS2* gene among patients infected with COVID-19 [23]. In Italian population, a lower frequency of the allele variant allele T was observed, compared to data obtained from the Europeans [26]. The presence of the T allele is associated with a decrease in severe cases of COVID-19 [39, 43]. This substitution of G>T may lead to a decrease in the ability of *TMPRSS2* to bind directly to the S protein of SARS-CoV-2. This is due to the substitution of glycine for the long-side-chain hydrophobic valine, which causes a decrease in the functional activity of the proteinase and a reduction in the flexibility of the peptide due to the increase in hydrophobicity. However, the results obtained by Minashkin *et al.*, 2020, were discordant, indicating that individuals with the T allele are more susceptible to SARS-CoV-2 infection and have the highest risk of developing the severe form [40]. SNPS rs2298659, rs200291871, rs61735794, and rs61735792 were polymorphisms that were also associated with susceptibility, the development of the severe form of COVID-19, and the appearance of deaths caused by the disease. The study by Kim & Jeong (2021) was the only work that reported a strong correlation of the SNP rs2298659 (C, wild-type allele) with the increased case-fatality rate of COVID-19 [25]. Torre-Fuentes *et al.* (2021) identified that SNPs rs61735794 (A allele) and rs61735792 (T allele) were significantly associated with cases of SARS-CoV infection [23]. Analysis of the rs200291871 SNP in Italian patients with COVID showed the presence of risk allele C with a higher frequency than in other European countries [26].

Andolfo *et al.* (2021) analyzed the polymorphisms of chromosome 21 and found five SNPs within the *TMPRSS2* gene associated with protection against severe COVID-19, they are: rs3787946 (C allele, wild type), rs9983330 (G allele, wild type), rs12329760 (T allele), rs2298661 (A allele, wild type) and rs9985159 (T allele, wild type), and the last four also were significantly correlated with lower expression of *TMPRSS2* in lung tissues. Wild-type alleles of these SNPs were less detected in inpatients compared to healthy controls, indicating a protective role for these variants against disease progression [42].

Rokni *et al.* (2022) analyzed SNPs rs17859725 and rs4303795 in 288 hospitalized patients with COVID-19. The SNP rs17859725 (A>G, Ile256Ile) is a variant synonymous with the exchange of two isoleucine codons at position 256 [34]. In this SNP, the wild-type G allele was identified as associated with a marked increase in the risk of COVID-19 infection and the presence of the AG genotype was more frequent in patients who died [39]. rs4303795 is a functional SNP and is located 2 kb from the promoter region of the *TMPRSS2* gene. The results of allele frequencies detected in this SNP, the G allele, were predominantly found in hospitalized patients with COVID-19. Furthermore, the GG genotype was directly associated with increased development of severe form, as well as the occurrence of lung lesions [39].

Saih *et al.* (2021) used bioinformatics tools to understand the effects of mutations on the *TMPRSS2* protein. A total of six SNPs (rs781089181, rs570454392, rs867186402, rs12329760, rs118518290, rs762108701) were identified by the tools used as responsible for decreasing the stability of *TMPRSS2*, thus being considered harmful variants of the gene. The SNPs rs781089181, rs570454392, and rs867186402 proved to be important polymorphisms due to their location in crucial regions for gene function. The analysis of the native protein and its comparison with the mutant protein showed the formation of 5 disordered regions due the variants rs12329760 and rs118518290, which are responsible for altering the function of *TMPRSS2* [33].

From the results obtained, it is possible to have an overview of which and how the SNPs studied so far are related to the protection, susceptibility, and aggravation of COVID-19 cases. Analysis of these SNPs can help to understand the increase in cases and deaths in several countries where these studies were conducted and provides insights to future research in other countries where studies with SNPs have not yet been conducted. Additionally, the identification of these SNPs helps direct studies aimed at the development of therapeutic interventions that can help people with increased susceptibility or worse progression of the disease.

Conclusion

The *TMPRSS2* gene demonstrated several SNPs associated with the severity and clinical fate of patients infected with SARS-CoV-2. The SNP rs12329760 was the polymorphism most researched by several studies that associated the presence of the T allele with protection against COVID-19. In some SNPs, such as rs383510 and rs75603675, there was disagreement between the studies on which allele would be associated with susceptibility and/or severity of the disease. This divergence makes evident the need for further research to specify which allele is related to the worsening of cases. Furthermore, new studies that analyze *TMPRSS2* polymorphisms can improve the characterization of individuals with greater susceptibility, in the development of pharmaceutical therapies, and in the conception of new vaccines against COVID-19.

Application of research: Characterization of SNPs associated with COVID-19 severity. Determination of SNPs present in different populations and the evolution of cases.

Research Category: Immunogenetics; Infectious Diseases.

Abbreviations: SARS-CoV-2 - severe acute respiratory syndrome coronavirus 2; COVID-19 – Coronavirus disease 2019; CoVs-Coronaviruses; ARDS-acute respiratory distress syndrome; *TMPRSS2* -transmembrane serine protease 2; ETS-transcription factors transforming erythroblasts; CFR-case fatality rate CFR; eQTLs - quantitative expression traits loci; PubMed-National Library of Medicine National Institutes of Health of the USA; LILACS-Latin American and Caribbean Literature in Health Sciences; SCIELO-Scientific Electronic Library Online.

Acknowledgement / Funding: Authors are thankful to Postgraduate Program in Parasitic Biology in the Amazon (PPGBPA), University of State of Pará (UEPA), Brazil

****Principal Investigator or Chairperson of research: Caroliny Soares Silva**
University: University of State of Pará (UEPA), Brazil
Research project name or number: Review study

Author Contributions: CSS was responsible for the conceptualization, formal analysis, data curation, methodology, investigation, validation, visualization, roles/writing – original draft, editing. MJAS and DMS were responsible for the supervision, visualization, writing review and editing. KVBL and CCF were responsible for the supervision, visualization, writing review and editing. LNGCL performed the conceptualization, investigation, project administration, supervision, visualization, writing review and editing. All authors read and approved the final manuscript

Author statement: All authors read, reviewed, agreed and approved the final manuscript. Note-All authors agreed that- Written informed consent was obtained from all participants prior to publish / enrolment

Study area / Sample Collection: University of State of Pará (UEPA), Brazil

Cultivar / Variety / Breed name: Nil

Conflict of Interest: None declared

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.
Ethical Committee Approval Number: Nil

References

[1] De Vito A., Fiore V., Princic E., Geremia N., Panu Napodano C.M., Muredda A.A., Maida I., Madeddu G., Babudieri S. (2021) *PLoS One*, 16(3), 1-14.

[2] Guo Y.R., Cao Q.D., Hong Z.S., Tan Y.Y., Chen S.D., Jin H.J., Tan K.S., Wang D.Y., & Yan Y. (2020) *Military Medical Research*, 7(1), 1-10.

[3] Li J., Wang Y., Liu, Y., *et al.* (2022) *European Journal of Medical Research*, 27(1), 1-10.

[4] Dieter C., Brondani L.A., Leitão C.B., Gerchman F., Lemos N.E., Crispim D. (2022) *Plos One*, 17(7), 1-23.

[5] Lim S., Bae J.H., Kwon H.S. *et al.* (2021) *Nature Reviews Endocrinology*, 17, 11-30.

[6] Dhama K., Khan S., Tiwari R., Sircar S., Bhat S., Malik Y.S., Singh K.P., Chaicumpa W., Bonilla-Aldana D.K., & Rodriguez-Morales A.J. (2020) *Clin Microbiol Rev*, 33(4), 20-28.

[7] Astuti I. & Ysrafil. (2020) *Clinical Research and Reviews*, 14(4), 407-412.

[8] Rahman S., Shishir M. A., Hosen M. I., Khan M. J., Arefin A., & Khandaker A. M. (2022) *Gene Reports*, 27, 1-10.

[9] Glotov O. S., Chernov A. N., Scherbak S. G., & Baranov V. S. (2021) *Russian Journal of Genetics*, 57 (8), 878-892.

[10] Zhu N., Zhang D., Wang W., Li X., Yang B., Song J., Zhao X., Huang B., Shi W., Lu R., Niu P., Zhan F., Ma X., Wang D., Xu W., Wu G., Gao G.F., Tan W. (2020) *New England Journal of Medicine*, 382(8), 727-733.

[11] Fang L., Karakiulakis G., Roth M. (2020) *Lancet Respir Med.*, 8(4)

[12] [12] Hoffmann M., Kleine-Weber H., Schroeder S., Krüger N., Herrler T., Erichsen S., Schiergens T.S., Herrler G., Wu N.H., Nitsche A., Müller M.A., Drosten C., & Pöhlmann S. (2020) *Cell*, 181(2), 271-280.

[13] Lima L.N.G.C., Sousa M.S., Lima K.V.B. (2020) *J. Health Biol. Sci*, 8, 1.

[14] Zarubin A., Stepanov V., Markov A., Kolesnikov N., Marusin A., Khitrinskaya I., Swarovskaya M., Litvinov S., Ekomasova N., Dzhaubermezov M., Maksimova N., Sukhomyasova A., Shtygasheva O., Khusnutdinova E., Radzhabov M., Kharkov V. (2021) *Genes*, 12 (1), 1-16.

[15] Lam D.K., Dang D., Flynn A.N., Hardt M., Schmidt B.L. (2015) *Pain*, 156(5), 923-930.

[16] Thunders M. & Delahunt B. (2020) *Journal of Clinical Pathology*, 73(12), 773-776.

[17] Mollica V., Rizzo A., Massari F. (2020) *Future Oncology*, 16(27), 2029-2033

[18] Di Maria E., Latini A., Borgiani P., & Novelli G. (2020) *Human Genomics*, 14(1), 1-14.

[19] Ramos-Lopez O., Daimiel L., Ramirez de Molina A., Martínez-Urbistondo D., Vargas J.A., Martínez J.A. (2020) *International Journal of Genomics*, 1-8.

[20] Debnath M., Banerjee M., & Berk M. (2020) *FASEB Journal*, 34(7), 8787-8795.

[21] Teng S. & Tang Q. (2020) *Computational and Structural Biotechnology Journal*, 18, 2100-2106.

[22] Darbani B. (2020) *International Journal of Environmental Research and Public Health*, 17(10), 3433.

[23] Torre-Fuentes L., Matías-Guiu J., Hernández-Lorenzo L., Montero-Escribano, P., Pytel, V., Porta-Etessam, J., Gómez-Pinedo, U., & Matías-Guiu, J. A. (2021) Journal of Medical Virology, 93(2), 863-869.

[24] Asselta R., Paraboschi E.M., Mantovani A. & Duga S. (2020) *Aging*, 12(11), 10087-10098.

[25] Kim Y.C., & Jeong B.H. (2021) *Genes*, 12(1), 1-9.

[26] Latini A., Agolini E., Novelli A., Borgiani P., Giannini R., Gravina P., Smarrazzo A., Dauri M., Andreoni M., Rogliani P., Bernardini S., Helmer-Citterich M., Biancolella M., Novelli G. (2020) *Genes*, 11(9), 1-8.

[27] Irham L.M., Chou W.H., Calkins M.J., Adikusuma W., Hsieh S.L., Chang W.C. (2020) *Biochemical and Biophysical Research Communications*, 529(2), 263-269.

[28] Piva F., Sabanovic B., Cecati M., Giulietti M. (2021) *Eur J Clin Microbiol Infect Dis.*, 40(2), 451-455.

[29] Jeon S., Blazyte A., Yoon C., Ryu H., Jeon Y., Bhak Y., Bolser D., Manica A., Shin E.S., Cho Y.S., Kim B.C., Ryoo N., Choi H. & Bhak J. (2021) *Molecules and Cells*, 44(9), 680-687.

[30] Wulandari L., Hamidah B., Pakpahan C., Damayanti N.S., Kurniati N.D., Adiatmaja C.O., Wigianita M.R., Soedarsono, Husada D., Tinduh D., Prakoeswa C., Endaryanto A., Puspaningsih N., Mori Y., Lusida M.I., Shimizu K., & Oceandy D. (2020) *Human Genomics*, 15 (1), 1-9.

[31] Monticelli M., Hay Mele B., Benetti E., Fallerini C., Baldassarri M., Furini S., Frullanti E., Mari F., Andreotti G., Cubellis M.V., Renieri A. (2021) *Genes*, 12(4), 1-15.

[32] Senapati S., Kumar S., Singh A.K., Banerjee P., & Bhagavatula S. (2020) *Journal of Genetics*, 99(1), 1-5.

[33] Saih A., Bouqdayr M., Baba H., Hamdi S., Moussamih S., Bennani H., Saile R., Wakrim L., & Kettani A. (2021) *Biomed Res Int*, 9982729, 1-17.

[34] Vargas-Alarcón G., Posadas-Sánchez R., & Ramírez-Bello J. (2020) *Life Sci.*, 260 (118313), 1-13.

[35] Paniri A., Hosseini M.M., & Akhavan-Niaki H. (2020) *Journal of Biomolecular Structure and Dynamics*, 1-18.

[36] Schönfelder K., Breuckmann, K., Elsner C., Dittmer U., Fistera D., Herbstreit F., Risse J., Schmidt K., Sutharsan S., Taube C., Jöckel K.H., Siffert W., Kribben A., & Möhlendick B. (2021) *Frontiers in Genetics*, 12, 667231.

[37] Monticelli M., Mele B.H., Andreotti G., Cubellis M.V., & Riccio G. (2021) *European Journal of Medical Genetics*, 64 (6), 104227.

[38] Pandey R.K., Srivastava A., Singh P.P., & Chaubey G. (2022) *Infection, Genetics and Evolution*, 98, 105206.

[39] Rokni M., Heidari Nia M., Sarhadi M., Mirinejad S., Sargazi S., Moudi M., Saravani R., Rahdar S., & Kargar M. (2022) *Applied Biochemistry and Biotechnology*, 194(8), 3507-3526

[40] Minashkin M.M., Grigortsevich N.Y., Kamaeva A.S., Barzanova V.V., Traspov A.A., Godkov M.A., Ageev F.A., Petrikov S.S., & Pozdnyakova N.V. (2022) *Biomedicines*, 10 (3), 549.

[41] Smatti M.K., Al-Sarraj Y.A., Albagha O., & Yassine H.M. (2020) *Frontiers in Genetics*, 11, 578523.

[42] Andolfo I., Russo R., Lasorsa V.A., Cantalupo S., Rosato B.E., Bonfiglio F., Frisso G., Abete P., Cassese G.M., Servillo G., Esposito G., Gentile I., Piscopo C., Villani R., Fiorentino G., Cerino P., Buonerba C., Pierri B., Zollo M., Iolascon A. (2021) *Science*, 2021; 24 (4), 102322.

[43] Villapalos-García G., Zubiaur P., Rivas-Durán R., Campos-Norte P., Arévalo-Román C., Fernández-Rico M., García-Fraile Fraile L., Fernández-Campos P., Soria-Chacartegui P., Fernández de Córdoba-Oñate S., Delgado-Wicke P., Fernández-Ruiz E., González-Álvaro I., Sanz J., Abad-Santos F. & de Los Santos I. (2022) *Life Science Alliance*, 5 (10)

[44] Huang C., Wang Y., Li X., Ren L., Zhao J., Hu Y., Zhang L., Fan G., Xu J., Gu X., Cheng Z., Yu T., Xia J., Wei Y., Wu W., Xie X., Yin W., Li H., Liu M., Xiao, Y. (2020) *The Lancet*, 395 (10223), 497-506.

[45] Ovsyannikova I.G., Haralambieva I.H., Crooke S.N., Poland G.A., & Kennedy R.B. (2020) *Immunological Reviews*, 296(1), 205-219.

[46] Senapati S., Banerjee P., Bhagavatula S., Kushwaha P.P., & Kumar S. (2021) *Journal of Genetics*, 100(1), 12.

[47] Beyerstedt S., Casaro E. B., & Rangel É. B. (2021) *European journal of clinical microbiology & infectious diseases, official publication of the European Society of Clinical Microbiology*, 40(5), 905-919.

[48] Cheng Z., Zhou J., To K.K., Chu H., Li C., Wang D., Yang D., Zheng S., Hao K., Bossé Y., Obeidat M., Brandsma C.A., Song Y.Q., Chen Y., Zheng B.J., Li L., & Yuen K.Y. (2015) *Journal of Infectious Diseases*, 212 (8), 1214-1221.

[49] Bizzotto J., Sanchis P., Abbate M., Lage-Vickers S., Lavignolle R., Toro A., Olszevicki S., Sabater A., Cascardo F., Vazquez E., Cotignola J., & Gueron G. (2020) *Science*, 23(10), 101585.

ARTIGO 2 – Artigo submetido para publicação

Enfermedades Infecciosas y Microbiologia Clinica

Analysis of association of ACE2 and TMPRSS2 gene polymorphisms with COVID-19 in a cohort of professionals of the Brazilian Amazon

Análisis de la asociación de los polimorfismos de los genes ACE2 y TMPRSS2 con la COVID-19 en una cohorte de profesionales de la Amazonia brasileña

--Borrador del manuscrito--

Número del manuscrito:	
Tipo de artículo:	Original article / Original
Palabras clave:	SARS-CoV-2; Angiotensin-Converting Enzyme 2; TMPRSS2 SARS-CoV-2; Enzima Convertidora de Angiotensina 2; TMPRSS2
Autor correspondiente:	Marcos Jessé Abrahão Silva Universidade do Estado do Pará BRAZIL
Primer autor:	Caroliny Soares Silva
Orden de autores:	Caroliny Soares Silva
	Marcos Jessé Abrahão Silva
	Marcelo Cleyton da Silva Vieira
	Everaldina Cordeiro dos Santos
	Jeanne Gonçalves Cabral
	Alex Brito Souza
	Rebecca Lobato Marinho
	Sebastião Kauã de Sousa Bispo
	Karla Valéria Batista Lima
	Cristiane Cunha Frota
	Luana Nepomuceno Gondim Costa Lima
Resumen:	COVID-19 is an infectious disease caused by the SARS-CoV-2 virus, which can be symptomatic and have unfavorable outcomes. Its clinical progression can be the result of several factors, such as environmental, immunological, and genetic. In this sense, the COVID-19 pandemic has caused incalculable damage to humanity, and in its first wave, it particularly affected clinical and hospital workers worldwide. The objective of this study was to analyze the genotypic and allelic profile of SNPs in the ACE2 and TMPRSS2 genes that could act as a risk factor for COVID-19 in professionals who worked in healthcare institutions during the first wave of the pandemic. This is a case-control study with samples collected from professionals in 10 healthcare institutions of Belém, Pará. Epidemiological questionnaires and chest CT scans were analyzed to assess the cases. SNP frequencies were obtained through partial Sanger sequencing. The C/C genotype was found in 100% of the SNP rs961360700 of the ACE2 gene. For SNP rs2298659, 68.24% of the C/C genotype and 29.8% of the C/T genotype were detected, while only 1.29% of the T/T genotype were detected. The presence of this polymorphism was significant for the change in symptomatic phenotypes and the correlation between individuals without previous comorbidities and the onset of symptoms. These findings highlight the importance of investigating these polymorphisms in other populations, as well as analyzing other SNPs in these genes, which may help understand the worsening of COVID-19 cases. La COVID-19 es una enfermedad infecciosa causada por el virus SARS-CoV-2, que puede ser sintomática y tener resultados desfavorables. Su progresión clínica puede ser el resultado de varios factores, como ambientales, inmunológicos y genéticos. En este sentido, la pandemia de COVID-19 ha causado daños incalculables a la

	humanidad y, en su primera ola, afectó particularmente al personal clínico y hospitalario de todo el mundo. El objetivo de este estudio fue analizar el perfil genotípico y alélico de los SNP en los genes ACE2 y TMPRSS2 que podrían actuar como un factor de riesgo para la COVID-19 en profesionales que trabajaron en instituciones de salud durante la primera ola de la pandemia. Se trata de un estudio de casos y controles con muestras recolectadas de profesionales en 10 instituciones de salud de Belém, Pará. Se analizaron cuestionarios epidemiológicos y tomografías computarizadas de tórax para evaluar los casos. Las frecuencias de SNP se obtuvieron mediante secuenciación parcial de Sanger. El genotipo C/C se encontró en el 100% del SNP rs961360700 del gen ACE2. Para el SNP rs2298659, se detectó el 68,24 % del genotipo C/C y el 29,8 % del genotipo C/T, mientras que solo se detectó el 1,29 % del genotipo T/T. La presencia de este polimorfismo fue significativa para el cambio en los fenotipos sintomáticos y la correlación entre individuos sin comorbilidades previas y la aparición de síntomas. Estos hallazgos resaltan la importancia de investigar estos polimorfismos en otras poblaciones, así como de analizar otros SNP en estos genes, lo que podría ayudar a comprender el agravamiento de los casos de COVID-19.

Primera página: first page (incluye autores, filiación y correspondencia)

Analysis of association of *ACE2* and *TMPRSS2* gene polymorphisms with COVID-19 in a cohort of professionals of the Brazilian Amazon

Caroliny Soares Silva[1] , Marcos Jessé Abrahão Silva [1,*] , Marcelo Cleyton da Silva Vieira [1] , Everaldina Cordeiro dos Santos [1,2] ,Jeanne Gonçalves Cabral [2], Alex Brito Souza [1,2] , Rebecca Lobato Marinho [1] , Sebastião Kauã de Sousa Bispo [2], Karla Valéria Batista Lima [2] , Cristiane Cunha Frota [3] , Luana Nepomuceno Gondim Costa Lima [1,2]

[1] Postgraduate Program in Parasitic Biology in the Amazon (PPGBPA), University of State of Pará (UEPA), Belém 66087-670, PA, Brazil.

[2] Bacteriology and Mycology Section (SABMI), Evandro Chagas Institute (IEC), Ananindeua 67030-000, PA, Brazil.

[3] Federal University of Ceará, Fortaleza 60020-181, CE, Brazil.

***Corresponding author:** jesseabrahao10@gmail.com

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.

Author Contributions

Conceptualization, C.S.S., K.V.B.L. and L.N.G.C.L.; methodology, M.J.A.S., C.S.S. and L.N.G.C.L.; validation, A.B.S., K.V.B.L. and L.N.G.C.L.; formal analysis, C.S.S., M.J.A.S., E.C.S., J.G.C., A.B.S.; investigation, R.L.M., C.S.S., S.K.S.B., and J.G.C.; resources, J.G.C., K.V.B.L. and L.N.G.C.L., writing—original draft preparation, C.S.S., M.J.A.S., K.V.B.L., C.S.S., R.L.M. and A.B.S.; writing—review and editing, K.V.B.L. and L.N.G.C.L.; visualization, C.C.F., K.V.B.L. and L.N.G.C.L.; supervision, K.V.B.L. and L.N.G.C.L.; project administration, L.N.G.C.L.; funding acquisition, K.V.B.L. and L.N.G.C.L. All authors have read and agreed to the published version of the manuscript.

Funding

The author(s) declared that financial support was received for the research and publication of this paper. CSS was funded through Bolsa FAPESPA ICAAF N° 006/2018.

Clinical trial number

Not applicable.

Acknowledgments

The authors thank to all the participants enrolled in this study, thanks to the Molecular Biology Laboratory of Evandro Chagas Institute and to the State University of Pará.

Artificial Intelligence Use Statement

None AI software was used in this manuscript.

Ethical Considerations

This work was approved by the Research Ethics Committee of the State University of Pará—UEPA (CAAE: 38113620.1.0000.5174) and is related to the research project "Análise da resposta ao SARS-CoV-2 em relação aos achados radiológicos e/ou à susceptibilidade genética individual", with opinion number: 6.124.862. This research was carried out in accordance with the Helsinki Declaration and Resolution No. 466/2012 of the Brazilian National Health Council.

Informed Consent

Informed Consent Form was applied to all study participants.

Analysis of association of *ACE2* and *TMPRSS2* gene polymorphisms with COVID-19 in a cohort of professionals of the Brazilian Amazon

Análisis de la asociación de los polimorfismos de los genes *ACE2* y *TMPRSS2* con la COVID-19 en una cohorte de profesionales de la Amazonia brasileña

ABSTRACT

COVID-19 is an infectious disease caused by the SARS-CoV-2 virus, which can be symptomatic and have unfavorable outcomes. Its clinical progression can be the result of several factors, such as environmental, immunological, and genetic. In this sense, the COVID-19 pandemic has caused incalculable damage to humanity, and in its first wave, it particularly affected clinical and hospital workers worldwide. The objective of this study was to analyze the genotypic and allelic profile of SNPs in the ACE2 and TMPRSS2 genes that could act as a risk factor for COVID-19 in professionals who worked in healthcare institutions during the first wave of the pandemic. This is a case-control study with samples collected from professionals in 10 healthcare institutions of Belém, Pará. Epidemiological questionnaires and chest CT scans were analyzed to assess the cases. SNP frequencies were obtained through partial Sanger sequencing. The C/C genotype was found in 100% of the SNP rs961360700 of the ACE2 gene. For SNP rs2298659, 68.24% of the C/C genotype and 29.8% of the C/T genotype were detected, while only 1.29% of the T/T genotype were detected. The presence of this polymorphism was significant for the change in symptomatic phenotypes and the correlation between individuals without previous comorbidities and the onset of symptoms. These findings highlight the importance of investigating these polymorphisms in other populations, as well as analyzing other SNPs in these genes, which may help understand the worsening of COVID-19 cases.

Keywords: SARS-CoV-2; Angiotensin-Converting Enzyme 2; TMPRSS2.

RESUMEN

La COVID-19 es una enfermedad infecciosa causada por el virus SARS-CoV-2, que puede ser sintomática y tener resultados desfavorables. Su progresión clínica puede ser el resultado de varios factores, como ambientales, inmunológicos y genéticos. En este sentido, la pandemia de COVID-19 ha causado daños incalculables a la humanidad y, en su primera ola, afectó particularmente al personal clínico y hospitalario de todo el mundo. El objetivo de este estudio fue analizar el perfil genotípico y alélico de los SNP en los genes ACE2 y TMPRSS2 que

podrían actuar como un factor de riesgo para la COVID-19 en profesionales que trabajaron en instituciones de salud durante la primera ola de la pandemia. Se trata de un estudio de casos y controles con muestras recolectadas de profesionales en 10 instituciones de salud de Belém, Pará. Se analizaron cuestionarios epidemiológicos y tomografías computarizadas de tórax para evaluar los casos. Las frecuencias de SNP se obtuvieron mediante secuenciación parcial de Sanger. El genotipo C/C se encontró en el 100% del SNP rs961360700 del gen ACE2. Para el SNP rs2298659, se detectó el 68,24 % del genotipo C/C y el 29,8 % del genotipo C/T, mientras que solo se detectó el 1,29 % del genotipo T/T. La presencia de este polimorfismo fue significativa para el cambio en los fenotipos sintomáticos y la correlación entre individuos sin comorbilidades previas y la aparición de síntomas. Estos hallazgos resaltan la importancia de investigar estos polimorfismos en otras poblaciones, así como de analizar otros SNP en estos genes, lo que podría ayudar a comprender el agravamiento de los casos de COVID-19.

Palabras clave: SARS-CoV-2; Enzima Convertidora de Angiotensina 2; TMPRSS2.

INTRODUCTION

The disease caused by the coronavirus 2019 (COVID-19 - Coronavirus disease 2019) is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2 - Severe Acute Respiratory Syndrome Coronavirus 2), a virus discovered in late December 2019, which exhibited a high infectivity rate [1]. The pathogen primarily affects the respiratory system, causing colds although other organs may also be affected. The most common symptoms of COVID-19 include fever, cough, shortness of breath, and muscle pain [2].

The entry of SARS-CoV-2 into the cell occurs through its receptor, the angiotensin-converting enzyme 2 (ACE2), which is present in various organs such as the lungs, kidneys, heart, and gastrointestinal tract. The binding occurs through the receptor-binding domain of the SARS-CoV spike protein, which binds strongly to human ACE2. The proteases of the target cells activate the spike protein by cleaving it into S1 and S2 subunits. S2 undergoes a conformational change, including the insertion of FP into the target cell membrane and exposure of the HR1 domain. The human proteases that cleave and activate the spike protein in the process of fusion with the human membrane are transmembrane serine protease 2 (TMPRSS2) and the trypsin-like protease of the human airways (HAT) [3].

ACE2 is widely distributed in the human body, being attached to the membrane of cells in organ tissues such as the heart, kidneys, and to a lesser extent, in the lungs, where it is concentrated in type II alveolar cells. ACE2, a monocarboxypeptidase, is responsible for

degrading angiotensin II, a pro-hypertensive and pro-fibrotic peptide, to generate angiotensin 1-7, a vasodilator and cardioprotective product that plays a role in the regulation of the renin-angiotensin-aldosterone system (RAAS). From the alteration in fluid homeostasis or blood pressure, the activation of the RAAS begins through the release of renin into the blood, causing the conversion of angiotensinogen into angiotensin I (Ang I), through the angiotensin-converting enzyme (ACE), into Ang II. This mechanism is responsible for the RAAS and thus causes vasoconstriction and an increase in blood pressure [4].

The *ACE2* gene is located on the X chromosome, and one of the most studied polymorphisms is associated with hypertension and heart diseases, where the polymorphism in this region is linked to the worsening of COVID-19. This occurs due to the formation of the SARS-CoV-2/ACE2 complex that allows the virus to enter the host cell and causes a negative regulation on the surface of ACE2, resulting in a loss of cardioprotective function, and thus a decline in the cardiovascular system. The SNP rs961360700 (C>T) does not have many characterizations in the literature but is already described as a polymorphism that may hinder the direct interaction between ACE2 and S1, due to a lower binding affinity with the spike protein, potentially significantly reducing the entry of SARS-CoV-2 and thus the infection [5].

TMPRSS2 is a type II transmembrane protease expressed in epithelial cells of the respiratory, gastrointestinal, prostate tract, and other organs. It has been identified that TMPRSS2 serves as an entry gateway for SARS-CoV-2, in addition to aiding the entry of other viruses. Therefore, inhibiting TMPRSS2 activity becomes essential to inhibit viral action. The *TMPRSS2* gene is located on chromosome 21, specifically in the 21q22.3 region. The SNP rs2298659 (G>A, C) is the most found polymorphism. The isoform of this SNP would be caused by the presence of the A allele of the SNP, which would be responsible for creating a splicing protein that helps in the formation of the isoform, responsible for cleaving the S protein, facilitating the entry of SARS-CoV-2 [6].

The state of Pará, Northern Brazil has a total of 905,316 accumulated cases, with a total of 19,329 confirmed deaths, resulting in a mortality rate of 224.68/100 thousand inhabitants [7]. The healthcare professionals who worked during the first wave of the COVID-19 pandemic were directly and extremely exposed to high viral loads, thus presenting the highest risk of infection. Furthermore, during the initial phase of the pandemic, these professionals faced a greater risk of contamination due to the reuse and shortage of personal protective equipment (PPE) and the absence of defined decontamination and contingency protocols for SARS-CoV-2 [8]. This study sought to characterize how the genetic factors of this population and the

mutations present in essential genes for viral entry were determinants for illness and the worsening of the clinical condition of individuals on the first pandemic wave.

MATERIAL AND METHODS

Settings

This study uses a case-control approach, and it was authorized by the State University of Pará—UEPA Research Ethics Committee (CAAE: 38113620.1.0000.5174), with opinion number 6.124.862. In compliance with the Helsinki Declaration and Brazilian National Health Council Resolution No. 466/2012, this study was conducted [9,10]. It followed the recommendations of Strengthening the reporting of observational studies in epidemiology (STROBE) [11]. The study was carried out in 10 medical facilities in Belém-PA, Brazil (the Brazilian Amazon region). Convenience sampling was used to choose 214 professionals from the administrative, medical, and general services domains who were actively employed in medical centers that treated COVID-19 patients between April 1, 2021, and June 30, 2020. This cohort has already been used for epidemiological analysis under other circumstances (investigation of causal association for other different SNPs) by previous work by our research group [12–14].

According to the health vulnerability situation reported in some countries during the first pandemic wave, such as Brazil, all employees of Brazilian healthcare facilities during this time were exposed to SARS-CoV-2 due to a lack of established safety protocols, overcrowding in healthcare facilities, a lack of masks, and a failure to use the N95 mask [15]. The approach utilized to divide the cohort subjects into groups according to the parameters and case definition of COVID-19 analyzed, and presenting the variables used in the study were also previously reported in our other previously published study with other SNPs analyzed in the same cohort (Group AS= Asymptomatic Group and Group SI= Symptomatic Group, of which the symptomatic subjects were further subdivided by severity according to the classification generated by the CCT data —lung involvement in CCT>10%— into SCP [with lung involvement] and SSP [without lung involvement]) [12]. A questionnaire was carried out (Supplementary Material S1) and TCLE (Informed Consent Term) was used with the participants.

Laboratory procedures

Samples were gathered between January 6, 2021, and March 30, 2022. Blood samples obtained by venipuncture were kept at -20°C for use in later laboratory procedures at the Molecular Biology Laboratory – LABIMOL, Bacteriology and Mycology Section (SABMI) of the Evandro Chagas Institute (IEC). DNA extractions were performed using the Dneasy Blood & Tissue kit (QIAGEN®, Venlo, Netherlands), and the manufacturer's instructions were adhered to. For all these positions, the possible association between the *ACE2* and *TMPRSS2* SNPs and the severity and/or susceptibility to COVID-19 were evaluated.

The National Center for Biotechnology Information's (NCBI) dbSNP website (http://www.ncbi.nlm.nih.gov/snp/; viewed on October 12, 2020) provided information about SNPs (SNP identification, or ID) [16]. In order to type the SNPs rs961360700 of *ACE2* and , the corresponding genomic regions were deposited in GenBank and their primers for the Polymerase Chain Reaction (PCR) were created by the Primer3Plus v2.0 program (http://www.bioinformatics.nl/primer3plus/; accessed on April 20, 2020) [17]. PCR was carried out using the thermocycler and Platinum Taq DNA Polymerase, DNA-free (Invitrogen®, Thermo Fisher Scientific Corporation, Waltham, Massachusetts, USA). A 2% agarose gel containing 3.0 µL of the amplified results was electrophoresed using Sybr Safe (Invitrogen®, Thermo Fisher Scientific Corporation, Waltham, Massachusetts, USA) to view the amplified DNA fragments in a photodocumentation device.

The PCR products were purified in accordance with the manufacturer's instructions using the EasyPure PCR Purification Kit (TransGen Biotech Co.®, Beijing, Beijing, China). The DNA fragments were sequenced using the BigDye X-Terminator kit on an ABI 3130 Genetic Analyzer sequencer (Applied Biosystems®, Life Technologies, Thermo Fisher Scientific Corporation, Waltham, MA, USA). The Bioedit v7.2.5 program was used to visualize and analyze SNP areas of interest, and BLAST was then performed on the NCBI website (https://blast.ncbi.nlm.nih.gov/Blast.cgi/; accessed on October 01, 2024) [18].

Statistical analysis and presentation of data

The extracted data was organized in a Microsoft Office Excel spreadsheet. The post-hoc sampling power assessment of the sample N was performed using G*Power software version 3.1.9.7 with a goodness-of-fit Chi-square test, with an α error probability of 0.05 and an effect size of 0.3, considering the acceptability of sampling power from 0.8 [19]. The normality of the variables was assessed using the Kolmogorov-Smirnov test. For correlative analysis between dependent and independent variables, multinomial logistic regression was used, comparing each category against the reference category (Group AS) for both symptomatology

and severity parameters of COVID-19. The probability (p) ≤ 0.05 was considered statistically significant.

Regarding the genotyped data from the Belém cohort, genotype frequencies were tested for Hardy-Weinberg Equilibrium (HWE) using the chi-square ($\chi 2$) test with $p < 0.001$ as the cut-off point for the significance level through SNPs Analyzer v. 1.1.82. The frequencies of the SNP genotypes and alleles were analyzed together with the epidemiological data, also in a multinomial manner, to verify and measure the effects [20].

In silico analyzes of these *ACE2* and *TMPRSS2* SNPs

Several bioinformatics servers were used together to evaluate the roles of each of these SNPs in gene function and gene-protein interactions. SIFT enabled the analysis of their location and type, as well as the prediction of their effects (https://sift.bii.a-star.edu.sg/; accessed on June 2, 2025). Regarding the prediction of the functional impacts of these SNPs, PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/bgi.shtml; accessed on June 2, 2025), PhDSNP (https://snps.biofold.org/phd-snp/phd-snp.html; accessed on June 3, 2025), SNP&GO (https://snps.biofold.org/snps-and-go/snps-and-go.html; accessed on June 3, 2025), PredictSNP2 (https://loschmidt.chemi.muni.cz/predictsnp2/; accessed on June 3, 2025), CADD (https://cadd.gs.washington.edu/; accessed on June 2, 2025), DANN (https://cbcl.ics.uci.edu/public_data/DANN/; accessed on June 2, 2025), GWAVA (https://loschmidt.chemi.muni.cz/predictsnp2/; accessed on June 3, 2025), and FATHMM (https://fathmm.biocompute.org.uk/; accessed on June 3, 2025) were the nine research tools used in public genetic databases to assess deleterious or non-deleterious parameters.

Predictions of the type of protein structural and conformational changes resulting from the SNPs were made using PONDR-VLXT servers to assess the natural disorder of gene regions (https://www.pondr.com/; accessed on June 10, 2025), and I-Mutant2 (https://folding.biofold.org/i-mutant/i-mutant2.0.html; accessed on June 3, 2025) and MuPRO (https://mupro.proteomics.ics.uci.edu/; accessed on June 3, 2025) analyzed the free energy of binding and protein stability after mutation in the region. Furthermore, the HOPE Project (https://www3.cmbi.umcn.nl/hope/; accessed on June 10, 2025) was able to analyze the structural effects of a point mutation in a protein sequence, characterizing chemical modifications important to the protein domain.

RESULTS

Sampling power, normality of variables, Hardy-Weinberg Equilibrium (HWE) of individuals in this Belém labor cohort

The sampling of each SNP was analyzed to verify the sampling power of the obtained material. Therefore, the power for SNP rs961360700 of the *ACE2* gene was 0.92, while the power for SNP rs2298659 of the *TMPRSS2* gene was 0.94, both presenting a power level greater than 0.80. Furthermore, the Hardy-Weinberg equilibrium of the genes was calculated, where a p-value of 1 was found for the *ACE2* gene, showing that the null hypothesis was accepted, with no change in genotype constancy in the 212 participants analyzed. Regarding the *TMPRSS2* gene, a p-value of 0.33 was found, according to the analysis of 213 participants.

Epidemiology of the Belém labor cohort and COVID-19

The basic epidemiological characteristics of the individuals in this Belém labor cohort have already been described in previous studies by our research group under different genetic investigations [12–14]. Binary logistic regression was performed on the symptomatology data obtained from the analyzed groups, with a total of 42% (90/213) asymptomatic participants and 58% (123/213) symptomatic participants. Furthermore, the epidemiological characteristics of the studied cohort were evaluated. The results showed significance in the correlation between the presence of pre-existing comorbidities and the onset of the disease in the symptomatic form, as well as the correction of the asymptomatic form with the absence of comorbidities, with a value of $p<0.001$.

Table 1. Binary Logistic Regression for Epidemiological Characteristics and Symptoms of COVID-19.

Dependent Variables: Symptomatology (ASxSI)[a]						
Independent variables	**B**	**S.E.**	**Wald**	**df**	**Sig. (*p*-value)**	**Exp (B)**
Occupational Category			2,010	2	0,366	
Occupational Category (Administrative)	0,382	0,342	1,252	1	0,263	1,466

Occupational Category (Healthcare Professionals)	0,478	0,424	1,272	1	0,259	1,612
Age Range			2,815	2	0,245	
Age Range (19-34)	0,532	0,558	0,908	1	0,341	1,702
Age Range (35-50)	0,031	0,544	0,003	1	0,955	1,031
Sex (Female)	0,431	0,320	1,813	1	0,178	1,539
Associated Comorbidities (Pre-existing)	-1,216	0,360	11,399	1	**<0,001**	0,296
Constant	0,625	0,520	1,443	1	0,230	1,868

[a]Model Deviation= 16,47; R^2 de Cox & Snell= 0,074; R^2 Nagelkerke= 0,1; Exp (B) = Odds Ratio; df= degree of freedom.

Multinomial logistic regression analysis evaluated the association of epidemiological characteristics and the development of severe cases of COVID-19. It showed significant value when analyzing the association of the absence of comorbidities with symptomatic individuals who presented pulmonary involvement, with a p=0.01. Furthermore, significance was found with $p<0.001$ in the analysis of the correlation of the youngest age group of professionals with the appearance of symptoms, however without the presence of pulmonary involvement.

Table 2. Multinomial Logistic Regression for Epidemiological Characteristics and Severity of COVID-19.

Dependent Variables: Disease Severity Groups (ASxSCPxSSP)									
Independent Variables[a]		**B**	**Error**	**Wald**	**df**	**Sig. (p-value)**	**Exp(B)**	**95% Confidence Interval for Exp(B)**	
								Limite inferior	**Limite superior**
SCP	Intercept	0,299	0,707	,179	1	0,672			
	[Profession Category = Administrative]	-0,538	0,589	,835	1	0,361	0,584	0,184	1,852
	[Profession Category = General Services]	-0,502	0,719	0,488	1	0,485	0,605	0,148	2,477

	[Profession Category = Healthcare Professionals]	0[b]	.	.	0	.	.	.	.
	[Age Range = 19-34]	0,078	0,801	0,010	1	0,922	1,081	0,225	5,194
	[Age Range = 35-50]	-0,089	0,743	0,014	1	0,905	0,915	0,213	3,925
	[Age Range = >50]	0[b]	.	.	0	.	.	.	.
	[Sex = Male]	0,790	0,481	2,702	1	0,100	2,204	0,859	5,657
	[Sex = Female]	0[b]	.	.	0	.	.	.	.
	[Associated Comorbidities = No]	-2,164	0,483	20,053	1	**<0,001**	0,115	0,045	0,296
	[Associated Comorbidities = Yes]	0[b]	.	.	0	.	.	.	.
SSP	Intercept	-19,025	1,109	294,139	1	<0,001			
	[Profession Category = Administrative]	-0,241	0,858	0,079	1	0,779	0,786	0,146	4,224
	[Profession Category = General Services]	-17,089	4893,428	0,0001	1	0,997	3,788E-8	0,0001	.[c]
	[Profession Category = Healthcare Professionals]	0[b]	.	.	0	.	.	.	.
	[Age Range = 19-34]	16,626	0,755	485,327	1	**<0,001**	16625450,631	3787567,093	72977085,801
	[Age Range = 35-50]	16,632	0,0001	.	1	.	16721763,569	16721763,569	16721763,569
	[Age Range = >50]	0[b]	.	.	0	.	.	.	.
	[Sex = Male]	0,064	,862	0,006	1	0,941	1,066	0,197	5,781
	[Sex = Female]	0[b]	.	.	0	.	.	.	.
	[Associated Comorbidities = No]	0,263	1,128	0,054	1	0,816	1,301	0,143	11,865

	[Associated Comorbidities = Yes]	0[b]	.	.	0	.	.	.	.

[a] The reference category is: AS.

[b] This parameter is set to zero because it is redundant.

[c] A floating-point overflow occurred while calculating this statistic. Therefore, its value is set to system-missing.

[d] Model Tuning = 71,01; R^2 de Cox & Snell= 0,211; R^2 Nagelkerke= 0,267; R^2 McFadden= 0,152

Genotyping of ACE2 (rs961360700) and TMPRSS2 (rs2298659) SNPs from the Belém cohort and COVID-19

ACE2 SNP rs961360700 had 100% C/C genotype and 100% C alleles. *TMPRSS2* SNP rs2298659 had 68.24% C/C genotype and 29.8% C/T genotype and only 1.29% T/T. Regarding the analysis of symptom development, it was found that the presence of pre-existing comorbidities (p=0.005), as well as the presence of the *TMPRSS2* rs2298659 gene polymorphism with the exchange for C/T alleles (0.01) were determinants for the change in symptomatic phenotypes.

Table 3. Binary logistic regression of the relationship between COVID-19 symptoms and *TMPRSS2* SNP genotyping.

Dependent Variables: Symptomatology (ASxSI)[a]						
Independent Variables	**B**	**S.E.**	**Wald**	**df**	**Sig. (*p*-value)**	**Exp (B)**
Associated Comorbidities (Pre-existing)	-0,956	0,340	7,920	1	**0,005**	0,384
TMPRSS2 SNP rs2298659 (C/T + T/T)	-0,792	0,308	6,628	1	**0,01**	0,453
Constant	1,278	0,318	16,210	1	<0,001	3,591

[a]Deviation Model= 15,29; R^2 de Cox & Snell= 0,07; R^2 Nagelkerke= 0,094.

Multinomial logistic regression was performed to assess the relationship between increased severity of COVID-19 cases and genotyping of the rs2298659 SNP of the *TMPRSS2* gene. Significance was detected with a p=0.0001 in the correlation between individuals with the absence of previous comorbidities and the onset of symptoms. However, this test did not reveal a significant correlation between more severe cases and the presence of polymorphism in the analyzed gene.

Table 4. Multinomial logistic regression of the relationship between COVID-19 severity and *TMPRSS2* SNP genotyping.

Dependent Variables: Disease Severity Groups (ASxSPIxSWPI)									
Independent Variables[a]		**B**	**Error**	**Wald**	**df**	**Sig.**	**Exp(B)**	**95% Confidence Interval for Exp(B)**	
								Lower limit	**Upper limit**
SCP	Intercept	-0,188	0,455	0,170	1	0,680			
	[Associated comorbidities = No]	-2,171	0,463	21,978	1	**0,0001**	0,114	0,046	0,283
	[Associated comorbidities = Yes]	0[b]	.	.	0	.	.	.	.
	[*TMPRSS2* (C/T +T/T) = No]	0,807	0,497	2,630	1	0,105	2,240	0,845	5,937
	[*TMPRSS2* (C/T +T/T) = Yes]	0[b]	.	.	0	.	.	.	.
SSP	Intercept	-3,235	1,189	7,404	1	0,007			
	[Associated comorbidities = No]	0,384	1,108	0,120	1	0,729	1,469	0,168	12,874
	[Associated comorbidities = Yes]	0[b]	.	.	0	.	.	.	.
	[*TMPRSS2* (C/T +T/T) = No]	0,745	0,845	0,777	1	0,378	2,107	0,402	11,043
	[*TMPRSS2* (C/T +T/T) = Yes]	0[b]	.	.	0	.	.	.	.

a. The reference category is: AS.

b. b. This parameter is set to zero because it is redundant.

c. Final Model Fit = 23,757; Pseudo R quadrado Cox e Snell = 0,198; Nagelkerke= 0,250; McFadden= 0,140

Table 5 shows the comparison of allele profiles found in the Belém population and the database available in gnomAD v4.0.0. It was evidenced that the polymorphism rs961360700 was practically absent in our samples and is extremely rare in the gnomAD databases; its absence in the cohort is expected given the very low comparative minor allele frequency (MAF).

Meanwhile, for the SNP rs2298659 of the TMPRSS2 gene, the T frequency ≈ 0.163 observed in the cohort is consistent with the magnitude reported in databases and studies (a relatively common variant in various populations), although the exact frequency by subpopulation (Latino/Admixed American) varies, it is plausible that the cohort (Amazonian Brazil) presents T ~0.16 while gnomAD global/by subpopulation shows something between ~0.12–0.25 depending on the sample.

Table 5. Comparative minor allele frequency (MAF) analysis of SNPs rs961360700 and rs2298659 between the Belém cohort (this study) and the gnomAD v4.0.0 database.

SNP (gene) Allele of interest Frequency - cohort (Belém) — gnomAD (Global / reference in the literature)	SNP (gene) Allele of interest Frequency - cohort (Belém) — gnomAD (Global / reference in the literature)	SNP (gene) Allele of interest Frequency - cohort (Belém) — gnomAD (Global / reference in the literature)	SNP (gene) Allele of interest Frequency - cohort (Belém) — gnomAD (Global / reference in the literature)
rs961360700 (*ACE2*; C>T; D355N)	**T (alternative) / C (wild-type)**	**0.0 (100% C in the cohort; variant not detected).**	Extremely rare in gnomAD/exomes — reported in the literature as T ≈ 0.000012 (≈1.2×10⁻⁵) in gnomAD_exome databases (mainly observed in frequency in European studies published))[21].
rs2298659 (TMPRSS2; c.888C>T, synonymous G296G)	**T (alternative / C (wild-type)**	Genotypes: C/C 68.24%, C/T 29.80%, T/T 1.29% → allele C ≈ 0.8370, T ≈ 0.1630 (frequencies calculated from the reported genotypes).	Variable among populations; studies consulting gnomAD/databases report a global MAF around ~0.12–0.23 depending on the sample/article (e.g., values reported in different studies: ~0.121; there are also studies reporting ~0.23 for certain cohorts)[22].

In silico analysis of functional and conformational changes related to SNPs and their related proteins

In silico analysis showed a deleterious potential of SNP rs961360700 in the *ACE2* gene of approximately 70% (8/9) when analyzed by bioinformatics server tools. Negative binding energies (I-Mutant2 score = -0.86 and MuPRO score = -0.95) showed the presence of instability in the protein conformation after the presence of the SNP, which may cause damage to the protein, demonstrating a decrease in protein stability after the presence of the SNP. The same analysis was performed for the *TMPRSS2* gene to assess whether the SNP caused any deformation in the protein. However, all tools analyzed showed neutral results.

Table 6. *In silico* evaluation of function and stability of *ACE2* and *TMPRSS2* SNPs.

		Functional											Structural			
													I-Mutant 2		MuPRO	
Gene	Amino Acid	Type	SIFT	PolyPhen 2	PhD SNP	SNP&GO	PredictSNP2	FunSeq2	CADD	DANN	GWAVA	FATHMM	Stability	DDG value (kcal/mol)	Stability	DeltaG
ACE2	D355N	Non synonymous	[a]	Pro Damaging	Disease	Disease	Deleterious	Neutral	Deleterious	Deleterious	Deleterious	Deleterious	Decrease	-0.86	Decrease	-0.957
TMPRSS2	G296G	Synonymous	Tolerated	[a]	[a]	[a]	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	[a]	[a]	[a]	[a]

[a] Not found.

In silico analysis of the presence of the SNP in the *ACE2* gene showed that the presence of the polymorphism introduces a different amino acid into the transcript, causing changes in protein properties, which may disrupt the domain, change the functions performed by *ACE2* and abolish its function. Regarding conservation profile, the wild-type residue is very conserved, but a few other residue types have been observed at this position too. The mutation is located within a domain, annotated in UniProt as: Peptidase M2, Peptidyl-Dipeptidase (Figure 1).

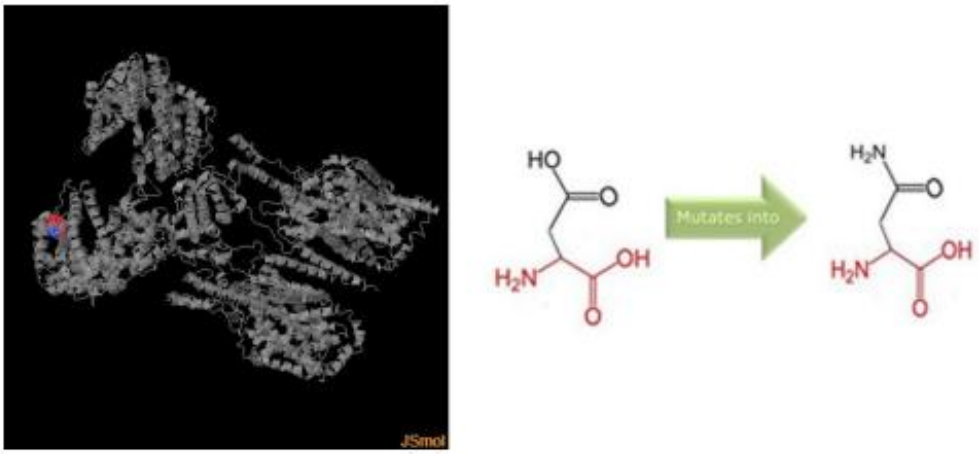


Figure 1. Protein model of *ACE2* after the presence of polymorphism and atomic representation of its modification after the presence of its SNP.

While the highlighted *TMPRSS2* SNP does not cause an amino acid change in the resulting protein structure. The graph of PONDR-VLXT analyzes the conformation of the resulting proteins and shows that the SNP of the *ACE*2 gene is in an area of order (D355N), where even with the presence of instabilities, the polymorphism would not cause significant changes in the protein order. However, the analysis of the *TMPRSS2* SNP shows that it is present in an area of disorder, where although it presents neutral potential, a functional dysfunction should occur with the presence of polymorphism, since it is in this area.

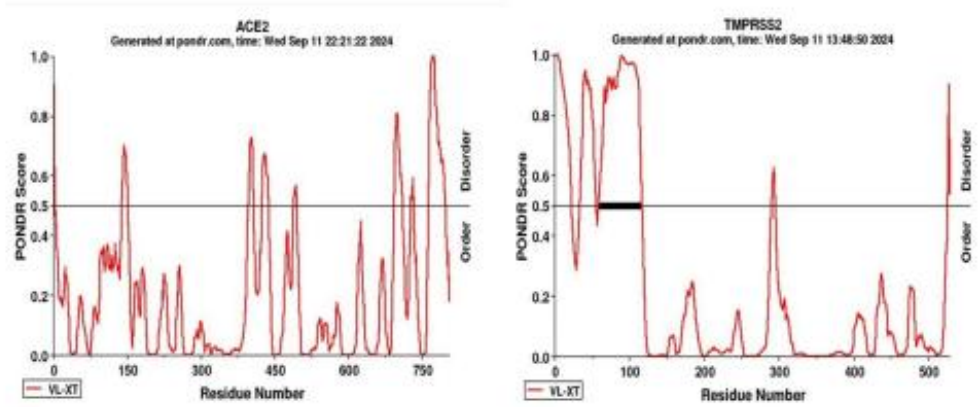


Figure 2. PONDR-VLXT Analyzes for Predicting Natural Disordered Regions of these SNPs regions.

DISCUSSION

The lack of testing, personal protective equipment, and health protocols has directly affected healthcare professionals in several countries. Studying genetic polymorphisms in the *ACE2* and *TMPRSS2* genes becomes important given the knowledge that genetic alterations can directly affect the entry of SARS-CoV-2 into host cells. Assessing this influence in individuals highly exposed to the virus, such as healthcare professionals in various fields, helps understand the transmission, symptoms, and progression of the disease.

The analysis of symptoms revealed that most participating professionals presented symptoms of the disease, demonstrating that contact with individuals infected with SARS-CoV-2 contributed to viral circulation among healthcare workers. These results corroborate previous studies showing that healthcare professionals who had contact with individuals infected or suspected of having COVID-19 had a higher chance of testing positively, a correlation with the onset of symptoms. Different groups of professionals were infected, regardless of whether they were on the frontline or not, highlighting the importance of universal testing for healthcare professionals [23].

Professionals with pre-existing comorbidities showed a significant correlation with the onset of symptomatic COVID-19. The detection of comorbidities was also detected in other studies, where hypertension, cardiovascular disease, or diabetes showed significant results in the study population, with the majority of patients receiving ICU care having hypertension. [24]. Additionally, hospitalized patients with comorbidities were more likely to experience severe or even fatal respiratory illnesses [25]. This worsening may be associated with low immune function, which may affect diabetic individuals and in cases of hypertensive patients.

It is important to emphasize that SARS-CoV-2 is responsible for positively regulating the expression of *ACE2* in hypertensive patients, which may cause an increase in blood pressure. This occurs because, in addition to playing an important role in viral entry, *ACE2* acts in the conversion of angiotensin 2 into angiotensin 1-7, participating in the renin-angiotensin-aldesterone system (RAAS), which regulates blood pressure [26]. Despite this, using multinomial logistic regression, another significant piece of data was also found showing that professionals without comorbidities were also symptomatic and presented pulmonary impairment, showing that other factors can also contribute to the worsening of cases.

In our study, younger professionals were the ones who most reported the onset of symptoms, but without lung involvement. Our cohort was considered younger when compared with other studies of healthcare professionals working at the beginning of the pandemic in European countries, where the majority of symptomatic participants were over 35 years old

[27,28]. This may be associated with the excessive increase in cases in the country where it was necessary to recruit young professionals to carry out emergency work during this period.

There are few studies that have analyzed the SNP rs961360700 of the *ACE2* gene, and even at the beginning of the pandemic it was described that this polymorphism was one of those responsible for the significant increase in binding free energy, which indicates a lower binding affinity, resulting in a lower susceptibility to SARS-CoV-2 infection. In our study, the presence of polymorphism was not detected. This may be associated with the low incidence of the SNP in the analyzed Brazilian Amazon population.

TMPRSS2 gene polymorphisms have been extensively studied in recent years, particularly regarding their association with COVID-19. In our study, analysis of the frequency of the rs2298659 SNP of the *TMPRSS2* gene showed that the majority had the wild-type C allele. From this, through binary logistic regression, it was possible to identify that, just as the presence of pre-existing comorbidities, the presence of the polymorphism affects the onset of symptoms in patients. It is known that disease progression can be affected by this polymorphism, impacting viral entry, disease severity, complications, and mortality. The polymorphism has previously been studied in American patients with severe disease, where its presence was significantly associated with a high risk of COVID-19 [6].

A comparison of the results obtained with the gnomAD data shows agreement between the polymorphism detection and the data presented worldwide, where the investigated *ACE2* gene polymorphism is rare in the global population, which corroborates the findings of our study. Regarding the SNP data of the *TMPRSS2* gene, a frequency within the expected global range was detected, demonstrating that the presence of the T allele behaved similarly in the Brazilian Amazon population.

In silico analysis of the rs961360700 SNP in the ACE2 gene allows us to visually demonstrate how genetic variants can modify the dynamics of the protein being affected. Our results demonstrate the deleterious potential of the analyzed SNP, with the presence of negative binding energies, causing instability in protein conformation. An in silico analysis study in the Indian population also showed that the same SNP presented a lower efficient binding energy (−895.2 kcal/mol) with tested SARS-CoV-2 variants that also had genomic alterations in their binding domain to the viral receptor, increasing individual susceptibility [29].

The construction of models for the simulation of molecular dynamics with SARS-CoV-2 showed an opposite effect, showing a significant increase in the free energy of binding, indicating a lower affinity and thus a reduction in viral infection [30]. The low sample size of this cohort and the low number of individuals among the participating disease severity groups may

have underestimated the statistical power of the significance of the data on greater morbidity in critical illness for COVID-19 in the presence of the *TMPRSS2* SNP.

As described here, the results predicted that despite the detection of negative binding energy, the presence of polymorphism can cause damage to the protein, where the presence of the SNP may possibly be associated with protection, preventing the binding of protein S to *ACE2*. Besides that, the PONDR-VLXT graphic shows that the SNP is in a region of instability. Furthermore, more studies *in vitro* and epidemiological ones with greater inferential statistical strength with sampling power must be carried out to confirm these data.

The rs2298659 SNP of the *TMPRSS2* gene is an intronic variant, meaning the polymorphism does not affect a coding region. *In silico* analysis showed that *TMPRSS2* remains neutral despite the presence of the SNP. Therefore, it is possible to correlate this with data obtained in the cohort of professionals studied, where the presence of the polymorphism did not present a protective effect, showing an association with symptomatic cases. Despite presenting a neutral result, the SNP is in a disordered area of the gene and can be considered one of the causes of its functional dysfunction. *In silico* analysis showed that the polymorphism impacts mRNA, affecting slicing and possibly generating more than 20 *TMPRSS2* isoforms [6].

CONCLUSIONS

Our study is the first to analyze these polymorphisms in Brazil and in healthcare professionals. The results showed that the polymorphism in the *TMPRSS2* gene was directly associated with the onset of symptoms in infected professionals. Furthermore, it was possible to verify that most patients who had contact with infected individuals presented symptoms, and that those with comorbidities were more likely to develop pulmonary involvement.

Few studies have been conducted with the SNPs investigated here. Our study highlights the importance of investigating these polymorphisms in other populations, aiming to confirm the *in silico* data obtained for *ACE2* rs961360700 regarding the deleterious role of the polymorphism, as well as to evaluate the data evidenced in our study, such as the role of the *TMPRSS2* gene rs2298659 polymorphism in the worsening of COVID-19.

REFERENCES

1. Senevirathne TH, Wekking D, Swain JWR, Solinas C, De Silva P. COVID-19: From emerging variants to vaccination. Cytokine & Growth Factor Reviews. 1º de abril de 2024;76:127–41.

2. Giuseppe Lippi FSG, Brandon M. Henry. COVID-19 and its long-term sequelae: what do we know in 2023? [Internet]. [citado 31 de maio de 2025]. Disponível em: https://www.mp.pl/paim/issue/article/16402

3. Lima LNGC, Sousa MS de, Lima KVB. As descobertas genômicas do SARS-CoV-2 e suas implicações na pandemia de COVID-19. Journal of Health & Biological Sciences. 15 de maio de 2020;8(1):1–9.

4. Pannucci P, Jefferson SR, Hampshire J, Cooper SL, Hill SJ, Woolard J. COVID-19-Induced Myocarditis: Pathophysiological Roles of ACE2 and Toll-like Receptors. Int J Mol Sci. 11 de março de 2023;24(6):5374.

5. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. JAMA. 7 de fevereiro de 2020;323(11):1061.

6. Posadas-Sánchez R, Fragoso JM, Sánchez-Muñoz F, Rojas-Velasco G, Ramírez-Bello J, López-Reyes A, et al. Association of the Transmembrane Serine Protease-2 (TMPRSS2) Polymorphisms with COVID-19. Viruses. 7 de setembro de 2022;14(9):1976.

7. Ministério da Saúde B. Coronavírus Brasil [Internet]. Coronavírus Brasil. 2025 [citado 5 de agosto de 2025]. Disponível em: https://covid.saude.gov.br/

8. Black JRM, Bailey C, Przewrocka J, Dijkstra KK, Swanton C. COVID-19: the case for health-care worker screening to prevent hospital transmission. Lancet. 2020;395(10234):1418–20.

9. Helsinki Declaration. World Medical Association's Declaration of Helsinki (with later amendments): Ethical principles for medical research involving human subjects. World Medical Association Helsinki; 1964.

10. Brasil MS. Resolução 466/12 do Conselho Nacional de Saúde. Sobre Diretrizes e Normas Regulamentadoras de Pesquisa envolvendo seres humanos. 2012;

11. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. International journal of surgery. 2014;12(12):1495–9.

12. Marcelino B dos R, Vieira MC da S, Silva MJA, Silva LCSS da, Gurrão EP da C, Santos EC dos, et al. Study of *TNF*-α, *IFN*-γ, *IL*-10, *TGF*-β and *IL*-6 Gene Polymorphisms in a Cohort of Professionals Who Worked in the First Pandemic Wave in the Brazilian Amazon. CRI [Internet]. 2025 [citado 16 de fevereiro de 2025];45(2). Disponível em: https://www.dl.begellhouse.com/journals/2ff21abf44b19838,444f93487bf4d1c4,2b2f3e00414789d0.html

13. Silva MJA, Silva CS, Marinho RL, Cabral JG, Gurrão EPDC, Dos Santos PAS, et al. Analysis of Epidemiological Factors and SNP rs3804100 of TLR2 for COVID-19 in a Cohort of Professionals Who Worked in the First Pandemic Wave in Belém-PA, Brazil. Genes. 5 de outubro de 2023;14(10):1907.

14. Silva MJA, Vieira MC da S, Souza AB, dos Santos EC, Marcelino B dos R, Casseb SMM, et al. Analysis of associations between the TLR3 SNPs rs3775291 and rs3775290 and COVID-19 in a cohort of professionals of Belém-PA, Brazil. Front Cell Infect Microbiol. 19 de dezembro de 2023;13:1320701.

15. Delgado D, Wyss Quintana F, Perez G, Sosa Liprandi A, Ponte-Negretti C, Mendoza I, et al. Personal safety during the COVID-19 pandemic: realities and perspectives of healthcare workers in Latin America. International Journal of Environmental Research and Public Health. 2020;17(8):2798.

16. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the National Center for Biotechnology Information in 2023. Nucleic acids research. 2023;51(D1):D29.

17. Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Research. 8 de maio de 2007;35(Web Server):W71–4.

18. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Em: Nucleic acids symposium series. [London]: Information Retrieval Ltd., c1979-c2000.; 1999. p. 95–8.

19. Kang H. Sample size determination and power analysis using the G*Power software. J Educ Eval Health Prof. 30 de julho de 2021;18:17.

20. Chen B, Cole JW, Grond-Ginsbach C. Departure from Hardy Weinberg Equilibrium and Genotyping Error. Frontiers in Genetics [Internet]. 2017 [citado 29 de junho de 2023];8. Disponível em: https://www.frontiersin.org/articles/10.3389/fgene.2017.00167

21. Cruz J de O, Conceição IMCA, Sousa SMB, Luizon MR. Functional prediction and frequency of coding variants in human ACE2 at binding sites with SARS-CoV-2 spike protein on different populations. J Med Virol. janeiro de 2021;93(1):71–3.

22. Monticelli M, Hay Mele B, Benetti E, Fallerini C, Baldassarri M, Furini S, et al. Protective Role of a TMPRSS2 Variant on Severe COVID-19 Outcome in Young Males and Elderly Women. Genes (Basel). 19 de abril de 2021;12(4):596.

23. Mani NS, Budak JZ, Lan KF, Bryson-Cahn C, Zelikoff A, Barker GEC, et al. Prevalence of Coronavirus Disease 2019 Infection and Outcomes Among Symptomatic Healthcare Workers in Seattle, Washington. Clin Infect Dis. 16 de junho de 2020;71(10):2702–7.

24. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395(10229):1054–62.

25. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020;395(10223):507–13.

26. Vosko I, Zirlik A, Bugger H. Impact of COVID-19 on Cardiovascular Disease. Viruses. 11 de fevereiro de 2023;15(2):508.

27. Van Loon N, Verbrugghe M, Cartuyvels R, Ramaekers D. Diagnosis of COVID-19 Based on Symptomatic Analysis of Hospital Healthcare Workers in Belgium. J Occup Environ Med. janeiro de 2021;63(1):27–31.

28. Havervall S, Rosell A, Phillipson M, Mangsbo SM, Nilsson P, Hober S, et al. Symptoms and Functional Impairment Assessed 8 Months After Mild COVID-19 Among Health Care Workers. JAMA. 18 de maio de 2021;325(19):2015–7.

29. Patil S, Alzahrani KJ, Banjer HJ, Halawani IF, Alzahrani H, Altayar MA, et al. Receptor binding domain of SARS-CoV-2 from Wuhan strain to Omicron B.1.1.529 attributes increased affinity to variable structures of human ACE2. J Infect Public Health. julho de 2022;15(7):781–7.

30. Wang J, Xu X, Zhou X, Chen P, Liang H, Li X, et al. Molecular simulation of SARS-CoV-2 spike protein binding to pangolin ACE2 or human ACE2 natural variants reveals altered susceptibility to infection. J Gen Virol. setembro de 2020;101(9):921–4.

7. CONCLUSÃO

- Esse estudo é o primeiro a analisar os polimorfismos rs*2298659* do gene *TMPRSS2* e o rs961360700 do gene *ACE-2* na população brasileira. A análise da frequência dos genótipos demostrou a prevalência do alelo selvagem no polimorfismo do gene *ACE-2* e a presença de mutações no SNP analisado do gene *TMPRSS2*.
- A análise das comorbidades existentes demostrou a presença de comorbidades pré existentes e o aparecimento da forma sintomática da covid-19.
- Também foi detectada uma relação significante entre indivíduos que não possuíam comorbidades com o aparecimento de comprometimento pulmonar em profissionais de saúde que apresentaram os sintomas;
- Quanto à genotipagem dos SNPs investigados, houve o aparecimento do alelo selvagem C no *rs961360700* do gene *ACE-2* em todas as amostras analisadas.
- No *rs2298659* do gene *TMPRSS2* foi detectada a presença do polimorfismo sendo esse associado diretamente com a presença de comorbidades pré-existentes e com o aparecimento da forma sintomática nos participantes.
- A análise *in sílico* demostrou a presença de instabilidade na proteína com a presença do polimorfismo *rs961360700* do ACE-2, demostrando um poder de causar deformação na molécula, podendo dificultar o ancoramento do SARS-CoV-2 na célula.
- O *rs2298659* do gene *TMPRSS2* foi detectado como uma mutação intronica que não causa modificações na região codificante, sendo possível correlacionar com os dados encontrados, onde o polimorfismo não apresentou efeito protetor, estando correlacionado com casos sintomáticos da COVID-19.
- Assim, enfatizamos a necessidade de estudos em mais SNPs dos genes *ACE-2* e *TMPRSS2* que estão diretamente envolvidos na entrada viral na célula do hospedeiro. Dessa forma será possível caracterizar totalmente essas regiões importantes que podem ser fortemente utilizadas como alvo para impedir a entrada viral.

REFERÊNCIAS

AISENBREY, Christopher; BECHINGER, Burkhard. Multimerization of the heptad repeat regions of the SARS-CoV 2 spike protein. **Biochimica et Biophysica Acta (BBA) - Biomembranes**, v. 1866, n. 2, p. 184259, 1 fev. 2024.

ALSHARIF, W.; QURASHI, A. Effectiveness of COVID-19 diagnosis and management tools: A review. **Radiography (London, England : 1995)**, v. 27, n. 2, p. 682–687, maio 2021.

ANTONY, Priya; VIJAYAN, Ranjit. Role of SARS-CoV-2 and ACE2 variations in COVID-19. **Biomedical Journal**, [S.L.], v. 44, n. 3, p. 235-244, jun. 2021. Elsevier BV. http://dx.doi.org/10.1016/j.bj.2021.04.006.

ASAKURA, Hidesaku; OGAWA, Haruhiko. COVID-19-associated coagulopathy and disseminated intravascular coagulation. **International Journal of Hematology**, v. 113, n. 1, p. 45–57, 2021.

ANGELI, Fabio; REBOLDI, Gianpaolo; TRAPASSO, Monica; ZAPPA, Martina; SPANEVELLO, Antonio; VERDECCHIA, Paolo. COVID-19, vaccines and deficiency of ACE2 and other angiotensinases. Closing the loop on the. **European Journal Of Internal Medicine**, [S.L.], v. 103, p. 23-28, set. 2022. Elsevier BV.

ASHOUR, Laith. Roles of the ACE/Ang II/AT1R pathway, cytokine release, and alteration of tight junctions in COVID-19 pathogenesis. **Tissue Barriers**, v. 11, n. 2, p. 2090792, *[S.d.]*.

ASSELTA, Rosanna *et al.* ACE2 and TMPRSS2 variants and expression as candidates to sex and country differences in COVID-19 severity in Italy. **Aging**, v. 12, n. 11, p. 10087–10098, 5 jun. 2020.

ASTUTI, Indwiani; YSRAFIL. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): An overview of viral structure and host response. **Diabetes & Metabolic Syndrome**, v. 14, n. 4, p. 407–412, 2020.

ATTIQ, Ali *et al.* Cytokine Storm-Induced Thyroid Dysfunction in COVID-19: Insights into Pathogenesis and Therapeutic Approaches. **Drug Design, Development and Therapy**, v. 18, p. 4215–4240, 20 set. 2024a.

ATTIQ, Ali *et al.* Cytokine Storm-Induced Thyroid Dysfunction in COVID-19: Insights into Pathogenesis and Therapeutic Approaches. **Drug Design, Development and Therapy**, v. 18, p. 4215–4240, 20 set. 2024b.

BELIK, Milja *et al.* Long-term COVID-19 vaccine- and Omicron infection-induced humoral and cell-mediated immunity. **Frontiers in Immunology**, v. 15, p. 1494432, 21 nov. 2024.

BICHARA, C. D. A.; AMORAS, E. S. G.; VAZ, G. L.; BICHARA, C. N. C.; AMARAL, I. P. C.; VALLINOTO, A. C. R. Persistence of Anti-SARS-CoV-2 IgM antibody up to 8 months post-COVID-19. *Case Reports in Clinical Medicine*, v. 10, p. 227-231, 2021. DOI: 10.4236/crcm.2021.109029.

BLACK, James R. M. *et al.* COVID-19: the case for health-care worker screening to prevent hospital transmission. **Lancet (London, England)**, v. 395, n. 10234, p. 1418–1420, 2020.

BOAVENTURA, Viviane S. *et al.* Syndromic Surveillance Using Structured Telehealth Data: Case Study of the First Wave of COVID-19 in Brazil. **JMIR Public Health and Surveillance**, v. 9, p. e40036, 24 jan. 2023.

BÖHM, Michael *et al.* Coronavirus Disease 2019 (COVID-19) and its implications for cardiovascular care: expert document from the German Cardiac Society and the World Heart Federation. **Clinical Research in Cardiology**, v. 109, n. 12, p. 1446–1459, 2020.

BRASIL, Ministério da Saúde. **Coronavírus Brasil**. *[S.l.]*: Ministério da Saúde, *[S.d.]*. Disponível em: <https://covid.saude.gov.br/>.

CHAUDHARY, Mayank. COVID-19 susceptibility: potential of ACE2 polymorphisms. **The Egyptian Journal of Medical Human Genetics**, v. 21, n. 1, p. 54, 2020.

CHEN, Ying *et al.* How does the SARS-CoV-2 reinfection rate change over time? The global evidence from systematic review and meta-analysis. **BMC Infectious Diseases**, v. 24, p. 339, 21 mar. 2024.

CHENG, Lijia *et al.* A review of current effective COVID-19 testing methods and quality control. **Archives of Microbiology**, v. 205, n. 6, p. 239, 17 maio 2023.

CHUNG, Yiu-Sing *et al.* Comprehensive Review of COVID-19: Epidemiology, Pathogenesis, Advancement in Diagnostic and Detection Techniques, and Post-Pandemic Treatment Strategies. **International Journal of Molecular Sciences**, v. 25, n. 15, p. 8155, 26 jul. 2024a.

CHUNG, Yiu-Sing *et al.* Comprehensive Review of COVID-19: Epidemiology, Pathogenesis, Advancement in Diagnostic and Detection Techniques, and Post-Pandemic Treatment Strategies. **International Journal of Molecular Sciences**, v. 25, n. 15, p. 8155, 26 jul. 2024b.

FANG, Enyue *et al.* Advances in COVID-19 mRNA vaccine development. **Signal Transduction and Targeted Therapy**, v. 7, p. 94, 23 mar. 2022.

FERNANDES, Queenie *et al.* Emerging COVID-19 variants and their impact on SARS-CoV-2 diagnosis, therapeutics and vaccines. **Annals of Medicine**, v. 54, n. 1, p. 524–540, *[S.d.]*.

GUSEV, Evgenii *et al.* SARS-CoV-2-Specific Immune Response and the Pathogenesis of COVID-19. **International Journal of Molecular Sciences**, v. 23, n. 3, p. 1716, 2 fev. 2022.

HADJ HASSINE, Ikbel. Covid-19 vaccines and variants of concern: A review. **Reviews in Medical Virology**, v. 32, n. 4, p. e2313, jul. 2022.

HOFFMANN, Markus *et al.* Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity.

EBioMedicine, v. 65, p. 103255, 4 mar. 2021.

HUSSEIN, Hosni A. M. *et al.* SARS-CoV-2 outbreak: role of viral proteins and genomic diversity in virus infection and COVID-19 progression. **Virology Journal**, v. 21, p. 75, 27 mar. 2024.

IRHAM, Lalu Muhammad *et al.* Genetic variants that influence SARS-CoV-2 receptor TMPRSS2 expression among population cohorts from multiple continents. **Biochemical and Biophysical Research Communications**, v. 529, n. 2, p. 263–269, 20 ago. 2020.

IWATA-YOSHIKAWA, Naoko *et al.* TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection. **Journal of Virology**, v. 93, n. 6, p. e01815-18, 15 mar. 2019.

KATIYAR, Harshita *et al.* SARS-CoV-2 Assembly: Gaining Infectivity and Beyond. **Viruses**, v. 16, n. 11, p. 1648, nov. 2024.

KOTSEV, Stanislav Vasilev *et al.* Hypotheses and facts for genetic factors related to severe COVID-19. **World Journal of Virology**, v. 10, n. 4, p. 137–155, 25 jul. 2021.

LAPORTE, Manon; NAESENS, Lieve. Airway proteases: an emerging drug target for influenza and other respiratory virus infections. **Current Opinion in Virology**, v. 24, p. 16–24, jun. 2017.

LIMA, Luana Nepomuceno Gondim Costa; SOUSA, Maisa Silva de; LIMA, Karla Valéria Batista. As descobertas genômicas do SARS-CoV-2 e suas implicações na pandemia de COVID-19. **Journal of Health & Biological Sciences**, v. 8, n. 1, p. 1–9, 15 maio 2020.

LIPPI, Giuseppe; SANCHIS-GOMAR, Fabian; HENRY, Brandon M. COVID-19 and its long-term sequelae: what do we know in 2023? **Polish Archives of Internal Medicine**, v. 133, n. 4, p. 16402, 19 abr. 2023.

LUCAS, J. M. *et al.* The androgen-regulated type II serine protease TMPRSS2 is differentially expressed and mislocalized in prostate adenocarcinoma. **The Journal of Pathology**, v. 215, n. 2, p. 118–125, jun. 2008.

LUNDBERG, Alexander L. *et al.* Surveillance Metrics and History of the COVID-19 Pandemic in Central Asia: Updated Epidemiological Assessment. **JMIR Public Health and Surveillance**, v. 10, p. e52318, 28 ago. 2024.

MAREI, Hany E.; ALTHANI, Asmaa; AFIFI, Nahla; POZZOLI, Giacomo; CACECI, Thomas; ANGELINI, Franco; CENCIARELLI, Carlo. Pandemic COVID-19 caused by SARS-CoV-2: genetic structure, vaccination, and therapeutic approaches. **Molecular Biology Reports**, [S.L.], v. 48, n. 9, p. 6513-6524, 16 ago. 2021. Springer Science and Business Media LLC. http://dx.doi.org/10.1007/s11033-021-06630-4.

MOHAMADIAN, Malihe *et al.* COVID-19: Virology, biology and novel laboratory diagnosis. **The Journal of Gene Medicine**, v. 23, n. 2, p. e3303, fev. 2021.

MONTICELLI, Maria *et al.* Protective Role of a TMPRSS2 Variant on Severe COVID-19 Outcome in Young Males and Elderly Women. **Genes**, v. 12, n. 4, p. 596, 19 abr. 2021.

MUSSABAY, Karakoz *et al.* Gut Microbiome and Cytokine Profiles in Post-COVID Syndrome. **Viruses**, v. 16, n. 5, p. 722, 2 maio 2024.

NAL, Béatrice *et al.* Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. **The Journal of General Virology**, v. 86, n. Pt 5, p. 1423–1434, maio 2005.

OCHANI, Rohan Kumar *et al.* COVID-19 pandemic: from origins to outcomes. A comprehensive review of viral pathogenesis, clinical manifestations, diagnostic evaluation, and management. *[S.d.]*.

PANNUCCI, Patrizia *et al.* COVID-19-Induced Myocarditis: Pathophysiological Roles of ACE2 and Toll-like Receptors. **International Journal of Molecular Sciences**, v. 24, n. 6, p. 5374, 11 mar. 2023.

PENHA, Diana *et al.* CO-RADS: Coronavirus Classification Review. **Journal of Clinical Imaging Science**, v. 11, p. 9, 2021.

PERICO, Luca; BENIGNI, Ariela; CASIRAGHI, Federica; NG, Lisa F. P.; RENIA, Laurent; REMUZZI, Giuseppe. Immunity, endothelial injury and complement-induced coagulopathy in COVID-19. **Nature Reviews Nephrology**, [S.L.], v. 17, n. 1, p. 46-64, 19 out. 2020. Springer Science and Business Media LLC. http://dx.doi.org/10.1038/s41581-020-00357-4.

POSADAS-SÁNCHEZ, Rosalinda *et al.* Association of the Transmembrane Serine Protease-2 (TMPRSS2) Polymorphisms with COVID-19. **Viruses**, v. 14, n. 9, p. 1976, 7 set. 2022.

PRINGLE, Kirsty G.; PHILP, Lisa K. Impact of ACE2 on the susceptibility and vulnerability to COVID-19. **Journal of Endocrinology**, v. 258, n. 3, 1 set. 2023.

RAI, Praveen *et al.* Detection technologies and recent developments in the diagnosis of COVID-19 infection. **Applied Microbiology and Biotechnology**, v. 105, n. 2, p. 441–455, 2021.

REN, Wenlin *et al.* Susceptibilities of Human ACE2 Genetic Variants in Coronavirus Infection. **Journal of Virology**, v. 96, n. 1, p. e01492-21, *[S.d.]*.

SANTOS, M. B. *et al.* Distinct Roles of Th17 and Th1 Cells in Inflammatory Responses Associated with the Presentation of Paucibacillary Leprosy and Leprosy Reactions. **Scandinavian Journal of Immunology**, v. 86, n. 1, p. 40–49, 2017.

SENAPATI, Sabyasachi *et al.* Assessment of risk conferred by coding and regulatory variations of TMPRSS2 and CD26 in susceptibility to SARS-CoV-2 infection in human. **Journal of Genetics**, v. 99, n. 1, p. 53, 2020.

SENEVIRATHNE, Thilini H. *et al.* COVID-19: From emerging variants to vaccination. **Cytokine & Growth Factor Reviews**, v. 76, p. 127–141, 1 abr. 2024.

SHARMA, Anshika; AHMAD FAROUK, Isra; LAL, Sunil Kumar. COVID-19: A Review on the Novel Coronavirus Disease Evolution, Transmission, Detection, Control and Prevention. **Viruses**, v. 13, n. 2, p. 202, 29 jan. 2021.

SHEIKHIAN, Farzaneh; MOFRAD, Sahar Sadeghi; TARASHI, Samira; JAJIN, Morteza Ghazanfari; SAKHAEE, Fatemeh; AHMADI, Iraj; ANVARI, Enayat; SHEIKHPOUR, Mojgan; FATEH, Abolfazl. The impact of ACE2 polymorphisms (rs1978124, rs2285666, and rs2074192) and ACE1 rs1799752 in the mortality rate of COVID-19 in different SARS-CoV-2 variants. **Human Genomics**, [S.L.], v. 17, n. 1, p. 1-23, 16 jun. 2023. Springer Science and Business Media LLC. http://dx.doi.org/10.1186/s40246-023-00501-8.

SILVA, Caroliny Soares *et al.* THE ASSOCIATION BETWEEN TMPRSS2 GENETIC POLYMORPHISMS AND THE SUSCEPTIBILITY AND SEVERITY OF COVID-19: A SYSTEMATIC REVIEW. **THE ASSOCIATION BETWEEN TMPRSS2 GENETIC POLYMORPHISMS AND THE SUSCEPTIBILITY AND SEVERITY OF COVID-19: A SYSTEMATIC REVIEW**, v. 14, n. 4, p. 6, 2022.

SILVA, Marcos Jessé Abrahão *et al.* Hyperinflammatory Response in COVID-19: A Systematic Review. **Viruses**, v. 15, n. 2, p. 553, 16 fev. 2023a.

SILVA, Marcos Jessé Abrahão *et al.* Analysis of associations between the TLR3 SNPs rs3775291 and rs3775290 and COVID-19 in a cohort of professionals of Belém-PA, Brazil. **Frontiers in Cellular and Infection Microbiology**, v. 13, p. 1320701, 19 dez. 2023b.

SONG, Meijuan *et al.* IFN-γ decreases PD-1 in T lymphocytes from convalescent COVID-19 patients via the AKT/GSK3β signaling pathway. **Scientific Reports**, v. 14, p. 5038, 29 fev. 2024.

SUBBARAO MALIREDDI, R. K.; SHARMA, Bhesh Raj; KANNEGANTI, Thirumala-Devi. Innate Immunity in Protection and Pathogenesis During Coronavirus Infections and COVID-19. **Annual review of immunology**, v. 42, n. 1, p. 615–645, jun. 2024.

ŠUŠAK, Božo *et al.* Excretion of SARS-CoV-2 RNA in feces has no prognostic benefit in the outcome of COVID-19: A clinical and immunological study. **Biomolecules and Biomedicine**, v. 24, n. 4, p. 1016–1027, 1 ago. 2024.

TAHER, Ibrahim *et al.* Relevance between COVID-19 and host genetics of immune response. **Saudi Journal of Biological Sciences**, v. 28, n. 11, p. 6645–6652, nov. 2021.

TEBHA, Sameer Saleem *et al.* Chapter One - SARS-CoV-2—Virus structure and life cycle. *In*: ZAIDI, Asiya Kamber (Org.). **Progress in Molecular Biology and Translational Science**. Immunobiology of COVID-19. *[S.l.]*: Academic Press, 2024. v. 202 p. 1–23.

TELLIER, Raymond. COVID-19: the case for aerosol transmission. **Interface Focus**, v. 12, n. 2, p. 20210072, *[S.d.]*.

TURČIĆ, Marijana *et al.* Interaction of HERVs with PAMPs in Dysregulation of Immune Response Cascade Upon SARS-CoV-2 Infections. **International Journal of Molecular Sciences**, v. 25, n. 24, p. 13360, 12 dez. 2024.

VARTAK, Abhishek; SUCHECK, Steven J. Recent Advances in Subunit Vaccine Carriers. **Vaccines**, v. 4, n. 2, p. 12, 19 abr. 2016.

VOSKO, Ivan; ZIRLIK, Andreas; BUGGER, Heiko. Impact of COVID-19 on Cardiovascular Disease. **Viruses**, v. 15, n. 2, p. 508, 11 fev. 2023.

WANG, Jingfang *et al.* Molecular simulation of SARS-CoV-2 spike protein binding to pangolin ACE2 or human ACE2 natural variants reveals altered susceptibility to infection. **The Journal of General Virology**, v. 101, n. 9, p. 921–924, set. 2020.

WETTSTEIN, Lukas; KIRCHHOFF, Frank; MÜNCH, Jan. The Transmembrane Protease TMPRSS2 as a Therapeutic Target for COVID-19 Treatment. **International Journal of Molecular Sciences**, v. 23, n. 3, p. 1351, 25 jan. 2022.

WHO. **COVID-19 Cases, World**. *[S.l.]*: WORLD HEALTH ORGANIZATION, *[S.d.]*. Disponível em: <https://data.who.int/dashboards/covid19/cases>.

WIEDMANN, Felix *et al.* SARS-CoV-2 ORF 3a-mediated currents are inhibited by antiarrhythmic drugs. **Europace**, v. 26, n. 10, p. euae252, 16 out. 2024.

WOOSTER, Luke *et al.* **Polymorphisms in the ACE2 Locus Associate with Severity of COVID-19 Infection**. medRxiv, , 22 jun. 2020. Disponível em: <https://www.medrxiv.org/content/10.1101/2020.06.18.20135152v1>. Acesso em: 30 set. 2025

XIANG, Qi *et al.* SARS-CoV-2 Membrane protein regulates the function of Spike by inhibiting its plasma membrane localization and enzymatic activity of Furin. **Microbiological Research**, v. 282, p. 127659, 1 maio 2024.

YADAV, Rohitash; CHAUDHARY, Jitendra Kumar; JAIN, Neeraj; CHAUDHARY, Pankaj Kumar; KHANRA, Supriya; DHAMIJA, Puneet; SHARMA, Ambika; KUMAR, Ashish; HANDU, Shailendra. Role of Structural and Non-Structural Proteins and Therapeutic Targets of SARS-CoV-2 for COVID-19. **Cells**, [S.L.], v. 10, n. 4, p. 821, 6 abr. 2021. MDPI AG. http://dx.doi.org/10.3390/cells10040821.YANG, Sijie *et al.* Fast evolution of SARS-CoV-2 BA.2.86 to JN.1 under heavy immune pressure. **The Lancet. Infectious Diseases**, v. 24, n. 2, p. e70–e72, fev. 2024.

YUAN, Yongliang *et al.* The development of COVID-19 treatment. **Frontiers in Immunology**, v. 14, p. 1125246, 26 jan. 2023.

YÜCE, Meral; FILIZTEKIN, Elif; ÖZKAYA, Korin Gasia. COVID-19 diagnosis —A review of current methods. **Biosensors & Bioelectronics**, v. 172, p. 112752, 15 jan. 2021.

ZHANG, Jin-jin *et al.* Risk and Protective Factors for COVID-19 Morbidity, Severity, and Mortality. **Clinical Reviews in Allergy & Immunology**, v. 64, n. 1, p. 90–107, 2023.

ZHAO, Tongxin *et al.* Protective RBD-dimer vaccines against SARS-CoV-2 and its variants produced in glycoengineered Pichia pastoris. **PLOS Pathogens**, v. 20, n. 8, p. e1012487, 30 ago. 2024.

ZLACKÁ, J; STEBELOVÁ, K; ZEMAN, M; HERICHOVÁ, I. Interactions of Renin-Angiotensin System and COVID-19: the importance of daily rhythms in ace2, adam17 and tmprss2 expression. **Physiological Research**, [S.L.], p. 177-194, 14 dez. 2021. Institute of Physiology of the Czech Academy of Sciences. http://dx.doi.org/10.33549/physiolres.934754.

APÊNDICE

APÊNDICE 1

genes

MDPI

Systematic Review

The Relationship between *TLR3* rs3775291 Polymorphism and Infectious Diseases: A Meta-Analysis of Case-Control Studies

Marcos Jessé Abrahão Silva [1], Caroliny Soares Silva [2], Marcelo Cleyton da Silva Vieira [2], Pabllo Antonny Silva dos Santos [2], Cristiane Cunha Frota [3], Karla Valéria Batista Lima [4] and Luana Nepomuceno Gondim Costa Lima [4,*]

1 Graduate Program in Epidemiology and Health Surveillance (PPGEVS), Evandro Chagas Institute (IEC), Ananindeua 67030-000, PA, Brazil; jesseabrahao10@gmail.com
2 Postgraduate Program in Parasitic Biology in the Amazon (PPGBPA), University of State of Pará (UEPA), Belém 66087-670, PA, Brazil; karolinysoares2303@gmail.com (C.S.S.); marcelocleiton14@hotmail.com (M.C.d.S.V.); antonnypabllo@gmail.com (P.A.S.d.S.)
3 Department of Pathology and Legal Medicine, Faculty of Medicine, Federal University of Ceará (UFC), Fortaleza 60441-750, CE, Brazil; cristianefrota71@gmail.com
4 Bacteriology and Mycology Section of the Evandro Chagas Institute (IEC), Ananindeua 67030-000, PA, Brazil; karlalima@iec.gov.br
* Correspondence: luanalima@iec.gov.br

Abstract: As the host's first line of defense against pathogens, *Toll-like receptors* (*TLRs*), such as the *TLR3*, are genes encoding transmembrane receptors of the same name. Depending on their expression, TLRs cause a pro- or anti-inflammatory response. The purpose of the article was to determine whether there is an association between the *Toll-like receptor 3* (*TLR3*) rs3775291 Single Nucleotide Polymorphism—SNP and susceptibility to infections. This review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines and was registered in PROSPERO under the code CRD42023429533. A systematic search for relevant studies was performed using PubMed, Scopus, SciELO, Google Scholar, and Science Direct by the MeSH descriptors and the Boolean Operator "AND": "Infections"; "TLR3"; "SNP", between January 2005 and July 2022. Summary odds ratios (ORs) and corresponding 95% confidence intervals (CIs) were calculated for genotypic comparison assuming a dominant genetic model (CT + TT vs. CC). A meta-analysis of 18 studies consisting of 3118 cases and 4368 controls found a significant association for risk between the presence of the *TLR3* SNP rs3775291 and infections as part of the general analysis (OR = 1.16, 95% CI = 1.04–1.28, $p = 0.004$). In the subgroups of continents, the SNP had a protective role in Europe for 1044 cases and 1471 controls (OR = 0.83, 95% CI = 0.70–0.99, $p = 0.04$); however, the Asian (for 1588 patients and 2306 controls) and American (for 486 patients and 591 controls) continents had an increase in infectious risk (OR = 1.37, 95% CI = 1.19–1.58, $p < 0.001$; OR = 1.42, 95% CI = 1.08–1.86, and $p = 0.01$, respectively). Heterogeneity between studies was detected ($I^2 = 58\%$) but was explained in meta-regression by the subgroup of continents itself and publication bias was not evident. The results of the meta-analysis suggest a significant association between the *TLR3* rs3775291 polymorphism and susceptibility to infections. Thus, when analyzing subgroups, the Asian and American continents showed that this SNP confers a higher risk against infections in a dominant genotypic model. Therefore, more studies are necessary to fully elucidate the role of *TLR3* rs3775291 in infections.

Keywords: *TLR3*; single nucleotide polymorphism; infectious diseases

Citation: Silva, M.J.A.; Silva, C.S.; da Silva Vieira, M.C.; dos Santos, P.A.S.; Frota, C.C.; Lima, K.V.B.; Lima, L.N.G.C. The Relationship between *TLR3* rs3775291 Polymorphism and Infectious Diseases: A Meta-Analysis of Case-Control Studies. *Genes* **2023**, *14*, 1311. https://doi.org/10.3390/genes14071311

Academic Editor: Domenico Lio

Received: 25 May 2023
Revised: 19 June 2023
Accepted: 20 June 2023
Published: 21 June 2023

1. Introduction

Since they are closely related to deprivation and unhealthy living conditions, infectious diseases pose a serious public health concern. In this sense, a pattern can be seen in the epidemiological indices of various diseases in relation to the development of the population

and the quality of life in a particular area [1]. In the world, they are the second leading cause of death [2].

There are certain characteristics of clinical presentation, such as symptomatic and asymptomatic patients, in an endemic community for a given disease. Immunological variability, which is mainly caused by changes in genetic background, is one of the causes of this, among other factors [3]. Therefore, to fully understand the various consequences of infections, either through susceptibility or protection against them, associative immunogenetic investigations are crucial [4].

Animals' innate immunity is dependent on pattern recognition receptors (PRRs), which are specialized in recognizing pathogen-associated molecular patterns (PAMPs) and then activating a signaling pathway to cause type I interferon (IFN-I)- and interleukin-1 (IL-1)-mediated pro-inflammatory reactions [5].

Toll-like receptors (*TLRs*) are transmembrane receptor genes that are PRRs that are found in endosomes or on the surfaces of immune cells [6]. These receptors mediate the production of cytokines necessary for effective immunization while also detecting pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs) in cells [7]. There can be several types of signaling pathways activated based on PAMP. Each member of this receptor family (TLR1-10) has different expression patterns in various organs and particular ligands to carry out this identification in humans [8]. TLRs are located mainly on the surfaces of macrophages and dendritic cells (DCs), among other cell types, as well as on the membranes of endosomes and lysosomes [9].

A double-stranded RNA (dsRNA)-TLR3 signaling complex, consisting of one dsRNA and two TLR3 molecules, is created when TLR3 binds to dsRNA. *TLR3* is a gene that encodes the protein of the same name with 904 amino acids, which is responsible for recognizing the dsRNA of infectious agents, a viral replication intermediate, in cellular endosomes. *TLR3* begins downstream signal transmission and induces the creation of the antiviral protein (AVP). This gene is located on the human chromosome 4q35.1 and has five exons (coding regions) [10].

Numerous epithelial cells, including fibroblasts, immune cells, neurocytes, and immune cells, carry TLR3, which is most widely expressed in the placenta and pancreas. TLR3 works through the TIR domain-containing adaptor-inducing IFN-β (TRIF)-dependent TLR signaling pathway and acts on DCs bearing antigens responsible for inducing antigen-specific immune responses mediated by lymphocytes [11].

Among the most varied types of polymorphism, single nucleotide polymorphism (SNP) is a punctual change of nucleotide that can occur in introns (noncoding regions) or exons [12]. In this sense, the *TLR3* SNP rs3775291 is a non-synonymous mutation (Cytosine to Thymine, C > T) of the missense type in exon 4, that is, it causes a change in the codon of amino acids from leucine (Leu) to phenylalanine (Phe) at residue 412 and its presence results in hypoactivity of the receptor in the human organism [13].

The level of *TLR3* transcript is not affected by this SNP, but it was discovered to decrease the ability of *TLR3* to attach to dsRNA [13]. The function of *TLR3* is only partially compromised by the substitution of Leu412Phe, resulting in an attenuated inflammatory reaction. The solenoid protein structure becomes unstable if 412Phe is present, and this alters any possible glycosylation of the nearby residue Asn413 (which was found to have N-acetylglucosamines attached). In this case, the ectodomain of the *TLR3* receptor is formed by the Leu412Phe variation, which is close to the glycosylation location (Asn413) and a crucial region for the dimerization of the domain at the membrane [14].

Investigating *TLR3* SNPs and diseases for the characterization of biomarkers in populations is of enormous immunogenetic relevance in this context. Using case-control epidemiological research, this work seeks to summarize and assess the relationship between the *TLR3* SNP rs3775291 and infectious illnesses. Because bias can occur in original studies as a result of flaws in the design of the included study, which tend to skew the magnitude or direction of associations in the data, the case-control study design was chosen as the selection strategy for the research added to this meta-analysis [15].

2. Material and Methods

2.1. Study Design

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [16]. To create the guiding question, the PICO strategy was used with the following anagrams: population, intervention, comparison, and outcome. In this context, it was developed from Population: patients with infectious diseases; Intervention: association between *TLR3* SNP rs3775291 and infectious diseases; Comparison: infectious diseases and *TLR3* SNP rs3775291; Outcome: identification of susceptibility or protective functions of *TLR3* SNP rs3775291 for infectious diseases published in the literature [17]. This review was registered in PROSPERO under the code CRD42023429533.

2.2. Search Strategy

The identification and selection of articles were performed in the databases Science Direct, the National Library of Medicine National Institutes of Health of the USA (PUBMED), Google Scholar, Scientific Electronic Library Online (SciELO), and Scopus using the descriptors: "*TLR3*", "Infections", and "Single Nucleotide Polymorphism", together with the Boolean operator "AND". The time cutoff was designated from the beginning of publications on the *TLR3* SNP rs3775291 (January 2005) until July 2022. The data were collected on 21 July 2022. The languages included in the study were limited to English, Portuguese, and Spanish.

The study titles and abstracts were examined and those that included *TLR3* polymorphisms and the probability of contracting an infectious disease were considered for a more thorough analysis. Electronic surveys were conducted from inception to 21 July 2020. Studies that examined *TLR3* polymorphisms and their link to noninfectious illness, as well as studies that were not published in English, were also eliminated. A study was considered eligible if it met all three of the following requirements: (i) it revealed an infectious disease outcome; (ii) it was carried out using a case-control design, where "cases" are people who have the disease outcome and "controls" are people in the healthy population who do not; (iii) it revealed genotype frequencies for *TLR3* rs3775291 (+1234C/T); (iv) the genotyping method by molecular biology.

2.3. Data Extraction

Two authors (MJAS and CSS) independently extracted all data that were considered relevant, including the differences and ambiguities found in the publications, and in cases of inconsistent selection, a third author (MCSV) participated in data selection. The information extracted included the name of the author, the origin of the population, the number of cases and controls subdivided by genotype frequencies (wild-type, heterozygous, and mutant homozygous), the disease being studied, and the conclusions reportedly drawn from each study.

Data extraction was conducted concurrently with an evaluation of the study's quality and bias by the ROBINS-E risk tool (Risk of Bias In Non-randomized Studies—of Exposures). Review Manager version 5.4 (Nordic Cochrane Centre, Cochrane Collaboration, Copenhagen, Sweden) was used to perform this analysis. Any disagreement between the two analyses of the researchers was discussed by a third author (MCSV) with regard to the inclusion of the study, any uncertainties in the data extraction process, and the quality and risk of bias evaluation within the larger group (MJAS and CSS).

2.4. Statistical Analysis

The Comprehensive Meta-Analyses—CMA program, version 2.2 (Biostat, Englewood, NJ, USA) was used on computer to perform the statistical analysis of the meta-analysis for the investigation. The fixed effects model estimated summary odds ratios (ORs) with 95% confidence intervals (95% CI). In each study's controls, the Hardy–Weinberg equilibrium (HWE) was evaluated using the chi-square (Q) goodness-of-fit test. Only genotypic com-

parison was performed using a dominant genetic model (genotypes CT and TT vs. CC). Using the I^2 statistics, the heterogeneity between studies was evaluated for comparisons.

According to Cochrane, the level of heterogeneity in a meta-analysis is analyzed based on the following parameters: 0% to 40%, it may not be important; 30% to 60% may represent moderate heterogeneity; 50% to 90% may represent substantial heterogeneity [18]. The chi-square test is one of the most commonly used tests to assess the significance of heterogeneity, with a significance level of $p < 0.05$ being used [19]. A subgroup analysis was performed according to the Continent. The Cochrane Q-test and I-squared (I^2) measure were used to determine the statistical difference groups ($p < 0.05$ was considered statistically significant). Begg's rank correlation test and a funnel plot were used to examine the potential for publication bias ($p < 0.05$ was considered statistically significant). Sensitivity analysis, meta-regression, and subgroup analysis based on study location were used to assess potential causes of variability, where applicable.

3. Results

3.1. Literature Search

Figure 1 summarizes the selection process. The search in four databases identified 169 articles, with only 32 nonduplicated or incomplete works. Subsequently, 95 articles were excluded based on titles and abstracts not relevant to the topic's theme or not associated with the searched SNP. Therefore, 42 articles were selected for complete reading, of which 12 were removed due to the type of study that did not correspond to the case-control. Regarding this number, 30 studies were eligible for inclusion; however, 12 were withdrawn due to incomplete data on the frequency of SNP allele or genotype in study subjects. Thus, this meta-analysis consisted of 18 case-control studies, mainly from PUBMED, and studies found in all languages covered by the methodology were included.

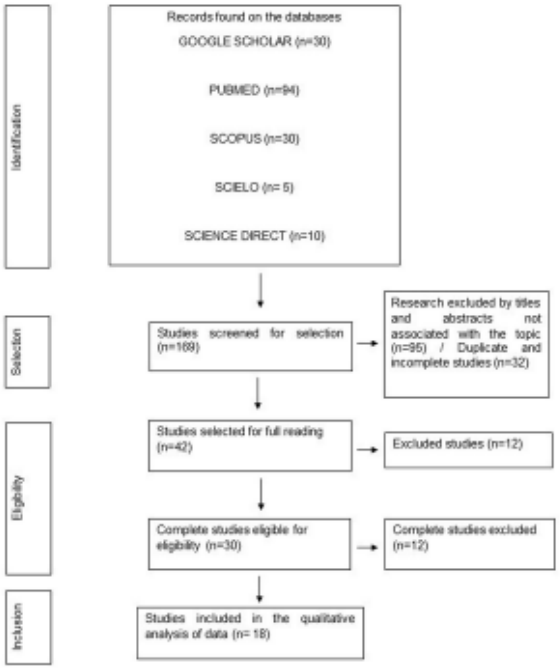


Figure 1. PRISMA flowchart representing the stages of selection, eligibility, and inclusion of studies for analysis. Belém, PA, Brazil (2022).

Figure 2 shows the composition of the articles included in terms of methodological quality assessment of the risk of bias in each. The evaluation was carried out separately by two evaluators in accordance with the uniform quality standard and was then cross-checked. When they ran into disagreements, they talked things out or asked the third author for guidance.

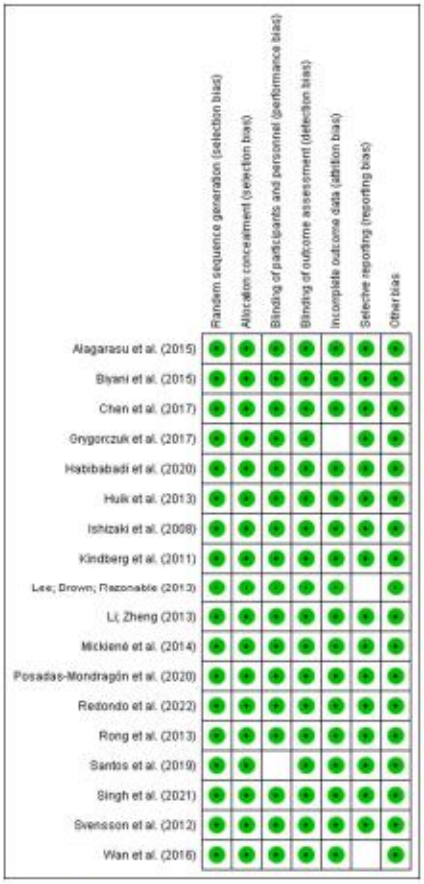


Figure 2. Risk of bias summary: review authors' judgments about each risk of bias item for each included study. Symbols in green mean compliance with the prerogative of that attribute investigated for that study, while the blank spaces (empty) demonstrate the gap for that information, and those in red indicate high methodological disagreement [13,20–36]. Source: Elaborated by the authors with RevMan v5.4 software.

3.2. Characteristics of the Included Studies

The basic characteristics of the studies, including the relationship between the *TLR3* rs3775291 polymorphism and the risk of infections, are found in Table 1. The data extracted came from surveys with populations from 12 countries on 3 continents (European, Asian, and American). The origin of the studies was in descending order, Asia (nine studies), Europe (six studies), and America (three studies). For the continent, most of the added studies were Asian (equivalent to 33.33%), while for the analysis towards countries, Chinese was the most added (n = 4, 22.22%) (Table 1).

Table 1. Characteristics of studies included in this review for *TLR3* SNP rs3775291.

Reference/Database	Type of Infectious Agent	Methodology/Genotyping Method/Population Size	Country/ Continent/ Ethnic Group	Gender Ratio (Male/Female)/Average Age of Participants	Absolute Count of Alleles and Genotypes (Cases/Controls)	HWE	*p*-Value	Results
Ishizaki et al. (2008) [20]/Science Direct	Virus/ssRNA-	Case-control/TaqMan essays/124 subjects (40 patients and 84 controls)	Japan/Asia/ Not reported.	Male: 27/Female: 13 and 84 children (with a no reported gender); Age: 12.4 years	T alleles (33/46) Dominant Model Genotypes (17/29)	$p > 0.05$	$p = 0.03$	The mutant T allele of the SNP has been associated with the risk of subacute sclerosing panencephalitis (SSPE) related to measles virus persistence.
Kindberg et al. (2011) [21]/PUBMED	Virus/ssRNA+	Case-control/PCR Genotyping/340 subjects (128 patients and 212 controls)	Lithuania/ Europe/Not reported.	Gender ratio and average age not reported.	T alleles (59/141) Dominant Model Genotypes (50/101)	$p > 0.05$	$p < 0.05$	The wild-type C allele has been associated with the risk of contracting tick-borne encephalitis virus (TBEV).
Svensson et al. (2012) [22]/PUBMED	Virus/dsDNA	Case-control/TaqMan Assay/401 individuals (239 patients and 162 controls)	Sweden/ Europe/Not reported.	Cases (Male: 136; Female: 103)/Controls (Male: 88; Female: 74); Age: 38 years.	T alleles (120/109) Dominant Model Genotypes (106/87)	$p > 0.05$	$p = 0.0272$	This SNP conferred protection against herpes simplex virus type 2 (HSV-2).
Lee; Brown; Razonable (2013) [23]/PUBMED	Virus/ssRNA+	Case-control/PCR Genotyping/611 subjects (153 patients and 458 controls)	USA/North America/Not reported.	Male: 395/Female: 216; Age: 52 years	T alleles (102/237) Dominant Model Genotypes (88/207)	$p > 0.05$	$p = 0.03$	The presence of the SNP promoted susceptibility to the Hepatitis C virus (HCV).
Li; Zheng (2013) [24]/PUBMED	Virus/dsDNA	Case-control/PCR Genotyping/948 subjects (466 patients and 482 controls)	China/Asia/Not reported.	Cases (Male: 383/Female: 83); Controls (Male: 386/Female: 96); Age: 53.8 years	T alleles (326/249) Dominant Model Genotypes (274/226)	$p > 0.05$	$p = 0.004$	It is a risk factor for developing HBV infection.
Rong et al. (2013) [25]/PUBMED	Virus/dsDNA	Case-control/PCR Genotyping/914 subjects (452 patients and 462 controls)	China/Asia/Not reported.	Cases (Male: 340/Female: 112)/Controls (Male: 344/Female: 118); Age: 38.88 years	T alleles (296/235) Model Genotypes (254/212)	$p > 0.05$	$p = 0.002$	The SNP increased the risk of HBV infection.

Table 1. *Cont.*

Reference/Database	Type of Infectious Agent	Methodology/Genotyping Method/Population Size	Country/ Continent/ Ethnic Group	Gender Ratio (Male/Female)/Average Age of Participants	Absolute Count of Alleles and Genotypes (Cases/Controls)	HWE	*p*-Value	Results
Huik et al. (2013) [26]/PUBMED	Virus/ssRNA+	Case-control/TaqMan Assay/842 subjects (172 HIV-positive patients and 670 controls)	Estonia/Europe/ White.	Cases (Male: 133/Female: 69)/Controls (Male:Not reported/Female: Not reported); Age: 30 years.	T alleles (108/455) Dominant Model Genotypes (92/376)	$p > 0.05$	$p = 0.03$	The SNP promoted HIV-1 protection.
Mickienė et al. (2014) [27]/PUBMED	Virus/ssRNA+	Case-control/PCR Genotyping/560 subjects (348 patients and 212 controls)	Lithuania/ Europe/ Not reported.	Cases (Male: 195/Female: 154); Controls (Male: Not reported/Female: Not reported). Age—Children TBE (Cases: 12.07/Controls: 11.43 years) Age—Adult severe TBE (Cases: 51.93 ± 15.419/Controls: 57.27 ± 15.108) Age—Adult TBE (Cases: 43.56/Controls: 46.97 years)	T alleles (195/141) Dominant Model Genotypes (157/101)	$p > 0.05$	$p = 0.02$	The SNP confers less risk of getting TBEV infection.
Alagarasu et al. (2015) [28]/PUBMED	Virus/ssRNA+	Case-control/PCR-RFLP/229 subjects (120 patients and 109 controls)	India/Asia/Not reported.	Cases (Male:73/Female: 47)/Controls (Male: 67/Female: 42); Age: 31.3 years	T alleles (59/63) Dominant Model Genotypes (52/48)	$p > 0.05$	$p = 0.04$	The SNP confers a greater risk of acquiring the dengue virus.

Table 1. *Cont.*

Reference/Database	Type of Infectious Agent	Methodology/Genotyping Method/Population Size	Country/ Continent/ Ethnic Group	Gender Ratio (Male/Female)/Average Age of Participants	Absolute Count of Alleles and Genotypes (Cases/Controls)	HWE	*p*-Value	Results
BiyaniByyani et al. (2015) [29]/PUBMED and Science Direct	Virus/ssRNA+	Case-control/PCR Genotyping/ 206 individuals (103 patients and 103 controls)	India/Asia/Not informed.	Gender ratio not reported; Age: 18.04.	T alleles (60/33) Dominant Model Genotypes (45/28)	$p > 0.05$	$p = 0.013$	The SNP has been associated with the risk of Japanese encephalitis virus (JEV) infection.
Wan et al. (2016) [30]/PUBMED	Virus/dsDNA	Case-control/TaqMan Assay/563 subjects (35 patients and 528 controls)	China/Asia/Not reported.	Cases (Male: 18/Female: 26)/Controls (Male: 358/Female: 295); Age: 26.43 years.	T alleles (17/329) Dominant Model Genotypes (17/282)	$p > 0.05$	$p = 0.736$	No significant association for SNP and HBV infection.
Chen et al. (2017) [31]/PUBMED	Virus/dsDNA	Case-control/PCR Genotyping/978 subjects (292 patients and 686 controls)	China/Asia/Not reported.	Male: 686/Female: 0; Age: 37 years.	T alleles (240/235) Dominant Model Genotypes (195/415)	$p > 0.05$	$p = 0.0001$	The SNP is a protective factor for HBV infection.
Grygorczuk et al. (2017) [32]/PUBMED	Virus/ssRNA+	Case-control/TaqMan Assay/180 subjects (108 patients and 72 controls)	Poland/ Europe/Not informed.	Gender: Not reported; Age: 42.44 years.	T alleles (59/41) Dominant Model Genotypes (51/36)	$p > 0.05$	$p < 0.05$	The higher frequency of wild-type C allele in patients was associated with TBEV infection.
Santos et al. (2019) [33]/PUBMED	Virus/ssRNA+	Case-control/Quantitative Real Time-PCR/ 255 subjects (168 patients and 87 controls)	Brazil/South America/Not reported.	Not reported.	T alleles (94/43) Dominant Model Genotypes (83/37)	$p > 0.05$	$p = 0.042$	The SNP increased risk of developing Zika virus infection.
Posadas-Mondragón et al. (2020) [34]/PUBMED	Virus/ssRNA+	Case-control/Real Time PCR/211 participants (165 patients and 46 controls)	Mexico/North America/Not reported.	DEN: 88/77; GP: Not reported; Age—Cases: 39.85, Controls: 49.52.	T alleles (96/56) Dominant Model Genotypes (83/23)	$p > 0.05$	$p = 0.31$	The mutant T allele of this SNP has been associated with protection from the dengue virus.
Habibabadi et al. (2020) [13]/PUBMED	Virus/ssRNA+	Case-control/PCR-RFLP/218 subjects (100 patients and 118 controls)	Iran/Asia/Not reported.	Male: 179/Female: 39; Age—Case Group: 38.55, Control Group: 36.72 years.	T alleles (53/70) Dominant Model Genotypes (48/63)	$p > 0.05$	$p = 0.46$	No association between this SNP and human T-cell lymphotropic virus type 1 (HTLV-1).

Table 1. *Cont.*

Reference/Database	Type of Infectious Agent	Methodology/Genotyping Method/Population Size	Country/ Continent/ Ethnic Group	Gender Ratio (Male/Female)/Average Age of Participants	Absolute Count of Alleles and Genotypes (Cases/Controls)	HWE	*p*-Value	Results
Singh et al. (2021) [35]/PUBMED	Virus/ssRNA+	Case-control/PCR Genotyping/ 196 individuals (98 patients and 98 controls)	India/Asia/Not informed.	Cases (Male: 36/Female: 62)/Controls (Male: 39; Female: 59); Age: 37.7 years	T alleles (52/28) Dominant Model Genotypes (42/25)	$p > 0.05$	$p = 0.10$	The presence of the mutant T allele of this SNP was associated with dengue susceptibility.
Redondo et al. (2022) [36]/PUBMED	Virus/dsDNA	Case-control/TaqMan Assay/204 individuals (50 cases and 154 controls)	Spain/ Europe/Not reported.	Male: 146/Female: 58; Age: 54.6 years	T alleles (39/90) Dominant Model Genotypes (28/75)	$p > 0.05$	$p = 0.029$	Homozygous carriers of the T minor allele (TT genotype) of the highlighted SNP had a twofold increased risk of BK polyomavirus viremia (BKPyV) after kidney transplantation.

3.3. Results of a Meta-Analysis and Publication Bias between the TLR3 rs3775291 Polymorphism and the Risk of Infections in Subgroups of Continents

Meta-analysis was performed using a fixed effect, and the overall effect estimate is plotted as a diamond. In a general analysis, a significant correlation was found directed toward the higher risk of infections in 3118 cases and 4368 controls (OR = 1.16, 95% CI = 1.04–1.28, $p = 0.004$). The subgroup analysis of the six studies conducted in European populations ($n = 1044$ cases and 1471 controls) and the general estimates indicate the association between protection from infections and the presence of the mutant allele of this SNP (OR = 0.83, 95% CI = 0.70–0.99, $p = 0.04$). Genotypic comparisons for the analyses of Asian (for 1588 patients and 2306 controls) and American (for 486 patients and 591 controls) subgroups were both statistically significant for the higher risk of these diseases investigated in the literature (OR = 1.37, 95% CI = 1.19–1.58, $p < 0.001$; OR = 1.42, 95% CI = 1.08–1.86, $p = 0.01$, respectively) (Figure 3).

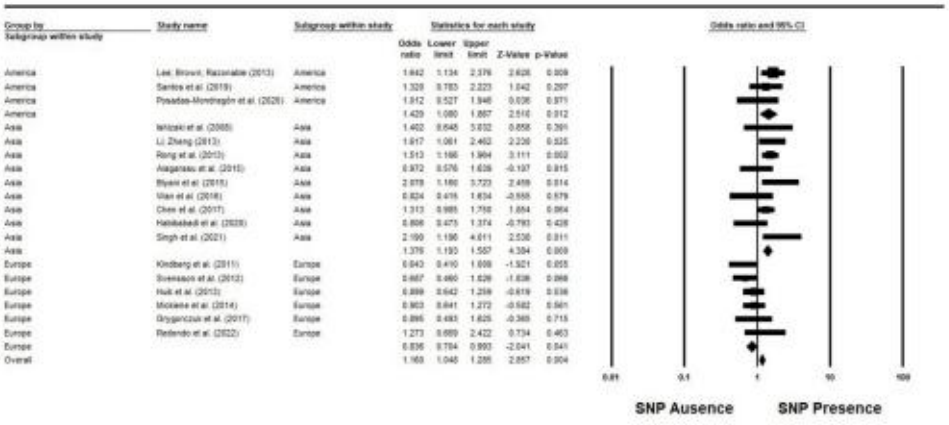


Figure 3. Forest plot of comparison about *TLR3* SNP rs3775291 and risk of infections, outcome: SNP presence for genotypes CT/TT vs. CC. The OR of each study is represented on the plot as a square with the area of each square proportional to the weight of the corresponding study in the meta−analysis. Horizontal lines are the 95% CIs associated for the OR of each study. The bold values highlight the total frequency of cases and controls, as well as the overall OR and the 95% CI [13,20–36]. Source: Elaborated by the authors with Comprehensive Meta−Analyses v2.2 software.

The standard error of the logarithm of the OR (SE(log[OR])) was plotted against the OR for each study. According to a widely accepted interpretation, when selection bias is present, the plot will become asymmetrical and the meta-analysis's overall impact will be skewed [20]. The symmetry, as in this study, in an inverted funnel shape implies the absence of publication bias (Figure 4).

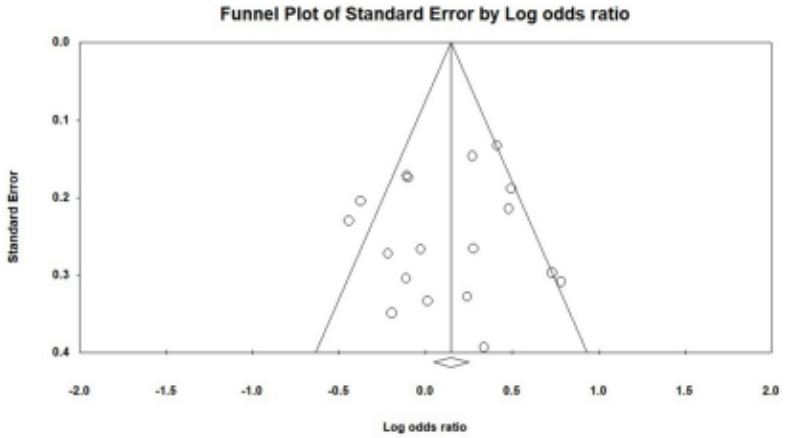


Figure 4. Funnel plot of comparison of the *TLR3* SNP rs3775291 and infection risk for all studies included in the meta-analysis, outcome: SNP presence for genotypes. Circles represent the included published studies and should be symmetrically dispersed around the overall effect in the form of an inverted funnel. Studies with higher precision are closer to the true value and situated at the narrowest part of the funnel. On the Y-axis of the graph, there is a measure of dispersion, the standard error, which is influenced by the sample size of the study. The larger this value is, the greater the inaccuracy of the study. On the X-axis of the graph, there is the effect measure measured in the meta-analysis and the center line is the result of this (which is directed on the X-axis by the diamond). The lines that make up the outline of the funnel correspond to the 95% CI. Source: Elaborated by the authors with Comprehensive Meta−Analyses v2.2 software.

3.4. Subgroup and Univariate Meta-Regression Analyses

Based on the generalized results found for this meta-analysis, a high level of heterogeneity was found between the included studies (Q = 40.58; $p < 0.001$; $I^2 = 58\%$). On the other hand, the investigation by subgroups based on continents revealed a low heterogeneity, referring to America (Q = 1.69; $p = 0.42$; $I^2 = 0\%$), Europe (Q = 4.29; $p = 0.50$; $I^2 = 0\%$), and Asia (Q = 13.07; $p = 0.10$; $I^2 = 38\%$). Despite this, a meta-regression analysis using a fixed model was conducted to confirm whether this factor was indeed the source of the observed heterogeneity. Figure 5 provides a visual observation of the Logit event rate by studying continent covariate (Q = 21.52; $p < 0.001$).

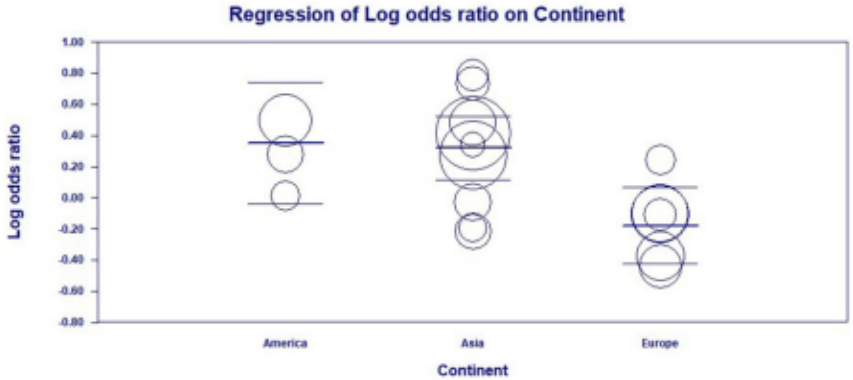


Figure 5. Bubble plot of the meta-regression analysis on the relationship between continent and the

risk of infection based on the presence of the rs3775291 SNP. The size of the bubble is inversely related to the variance of the study. The solid line represents the linear regression (continent as the meta-independent variable). The two lines in the horizontal area between the main line correspond to the confidence intervals of the prediction. Source: Elaborated by the authors with Comprehensive Meta−Analyses v2.2 software.

4. Discussion

TLR3 is a member of a family of immune receptors that are crucial for activating the innate immune response, indirect activation of the adaptive immune system, and control of cytokine expression in the defense of the body against infections [19,21,22]. TLR3 detects dsRNA, a molecular signature present in most viruses. TLR3 is critical for the induction of the antiviral state and the prevention of virus replication, but it can also promote an overactive and dysregulated immune response to infection, which is damaging to the host and helps to progress the severe form of the disease [23].

The presence of polymorphisms in *TLR3* is associated with changes in its structure and function, which can influence the immune response to viruses [24]. The TLR3–TRIF axis is essential in deciding how the balance between antiviral and immune regulatory pathways affects defensive versus offensive responses in the case of chronic viral infections of RNA that result in prolonged IFN-α/β signaling [25]. All articles on this *TLR3* SNP included in the meta-analysis were on viral infections. This is probably due to the binding of this receptor to its main ligand (dsRNA) in pathogens.

In this context, the molecular structure of a signaling unit (SU) reveals that dsRNA molecules attach to two regions: one near the N-terminus (LRR-NT and LRR1-LRR3) and one at the C-terminus (LRR19-LRR21). While protein–protein interactions only take place at LRR-CT in the TLR3-dsRNA complex, surface contacts (primarily hydrogen bonds and electrostatic interactions) are the only means by which TLR3 attaches to its receptors. To ensure TLR3 communication and dsRNA binding, the C-terminal dimerization region is essential [26].

The TRIF for downstream type I IFN signaling is shared by TLR3 and the DEAD (Asp-Glu-Ala-Asp) box polypeptide 1 (DDX1), DDX21, and DHX36 components of the DExD/H-box helicase cytosolic receptors of dsRNA. The only TLR receptor that relies exclusively on TRIF to trigger IFN-β release is TLR3. The TLR3-mediated signaling pathway can be divided into the TRIF-dependent nuclear transcription factor-κB (NF-κB) pathway and the TRIF-dependent IFN-regulatory factor 3/7 (IRF3/7) pathway based on the various downstream products that TRIF activates. For example, IRF3 induces the expression of type I interferons to mediate antiviral effects by activating other genes such as MxA genes [27].

TNF receptor-associated factor 3 (TRAF3) and TRAF6 interact with TRIF once it has been triggered in the plasmalemma by exogenous dsRNA to initiate a sequence of cascade events [28]. IRF3/7, NF-kB and activator protein 1 (AP-1) are the transcription factors that this signaling pathway eventually engages in, causing the production of type I IFNs (IFN-I), pro-inflammatory cytokines, and chemokines after TLR3 activation, respectively. TLR3 signaling also activates the proteins phosphoinositide 3-kinase (PI3K), p38-mitogen-activated protein kinase (MAPK), extracellular signal-regulated kinase (ERK), and c-Jun N-terminal kinase (JNK). Fine-tuning of ubiquitination and phosphorylation is essential for the TLR3 signaling cascade [29].

The meta-analysis carried out in this study explored the correlation between the rs3775291 polymorphism of the *TLR3* gene and the risk of infection and disease development in different countries. Analyzing the six studies carried out on the European continent, most of them (83.33%) have shown an association between the presence of the ancestral allele (C) and the risk of infection, while one of them detected that the mutant allele (T) may be involved with the risk of developing a disease, as seen in BKPyV infection after kidney transplantation [30]. Regarding the analysis of data from three American [31–33] and nine Asian studies [13,24,34–40], a significantly higher risk of infection with these diseases was found.

An essential line of defense in innate immunity related to the blood–brain barrier (BBB) is provided by a subpopulation of human neurons that expresses TLR3 on a constitutively. These hypotheses that TLR3 expression facilitates TBEV penetration through the BBB, which promotes the onset of neurologic illness, but also serves as a protective mechanism during established central nervous system (CNS) infection, may help explain these seemingly contradictory findings regarding the presence of this *TLR3* SNP [41]. Other investigations were conducted to examine the role of this SNP in responses to viral infections on the European continent [42,43]. These findings imply that in cases of TBE infection and HIV infection, the mutation may inhibit *TLR3* signaling activity, inhibiting viral entry into the central nervous system.

SNPs that cause loss of function of *TLR3*, such as rs3775291, appear to have an impact on the ability of the CNS to withstand HSV infection [44]. Studies in rodents show that astrocytes lacking TLR3 cannot produce an IFN-α response to HSV-2, predisposing these animals to an increased HSV-2 CNS infection if the peripheral (genital) infection does not worsen [45]. However, fibroblasts from TLR3-deficient patients with HSV-1 encephalitis (HSE) have abolished type I IFN activation, in contrast to PBMCs from the same patients, which respond properly to identical stimuli. Therefore, it is possible that the main function of TLR3 is to cause a type I IFN response to HSV-2 in non-hematopoietic cells rather than to support the acquired immune system [46].

According to a Spanish study, individuals who are homozygous carriers of the SNP T minor (TT) allele rs3775291 are two times more likely to develop BK polyomavirus (BKPyV) viremia [30]. TLR3 is involved in the stimulation of innate immune mechanisms during antiviral and inflammatory responses to BKPyV. Uncontrolled signaling caused by an SNP, such as rs3775291, can have an impact on the pathogenesis of BKPyV-associated nephropathy (BKPyVAN) by reducing its signaling activity compared to the wild-type type [25].

The study by Chen et al. (2017) makes additional assumptions that the missense polymorphism in rs3775291 may improve the anticancer immunostimulatory role (hepatocellular carcinoma—HCC linked to HBV) and support the apoptotic process [39].

The relationship between the *TLR3* rs3775291 polymorphism and HBV infection has been studied in Chinese populations [35,38]. Wan et al. (2016) were unable to find any evidence linking SNP to intrauterine HBV infection [38]. However, another study found that the mutant T allele was significantly more common in people with chronic hepatitis B and acute chronic renal failure, suggesting that this polymorphism may be a risk factor for the progression of the illness [35].

TLR3 has been associated with the release of cytokines and cellular activation caused by HCV [31]. On the other hand, the TLR3 ligand inhibits HBV proliferation in the liver of HBV mutant mice, according to recent investigations on infection [47]. The *TLR3* SNP rs3775291 in this situation impairs the secretion of the TLR3 ectodomain and reduces the receptor's ability to respond to antigenic stimuli from these viruses [48].

Arboviral outbreaks have a significant negative impact on the population of Latin America [49]. There have been investigations into the potential effects of the rs3775291 polymorphism on the clinical forms of arboviruses. Santos et al. (2019) examined the relationship between SNP and microcephaly caused by the Zika virus. The results of the study demonstrated a link between the existence of the T mutant allele in the SNP rs3775291 in the *TLR3* gene and the risk of congenital Zika syndrome (CZS) in pregnant women who contracted the virus during pregnancy. This link can be explained by the decline in the function of phenylalanine-containing *TLR3* (T allele), and it directly interferes with antiviral activity, leading to an increase in viral load and making it easier for the pathogen to enter the developing brain [32].

The Asian continent produced the majority of the studies found examining the relationship between the *TLR3* SNP rs3775291 and viral susceptibility and the progression of patient cases. An analysis of how this polymorphism affected dengue cases in India was performed. The study by Alagarasu et al. (2015) that involved the Indian population found

that the minor allele T has a negative impact on the presence of the protein structure and may be directly related to a reduction in inflammation, which would protect against the progression of dengue infections [36].

In addition to that, the measles virus (MV) appears to reach the CNS at the time of initial infection. Microglia and astrocytes, endogenous brain cells, are key players in the immunological responses that occur in this area. Type I IFN is induced by MV infection through TLR3, and in virus-infected cells, type I IFN is significantly upregulates TLR3 translation in a positive feedback way. Previous studies have shown that type I IFN inhibits efficient MV replication in rodents. Therefore, immunological dysregulation in measles is produced by uncontrolled signaling of TLR3 mediated by this polymorphism [24,50].

Furthermore, there are strong associations between this *TLR3* SNP and the differences in the downstream intracellular signaling molecules Myeloid Differentiation factor 2 (MD-2) and Myeloid Differentiation Primary Response 88 (MyD88) in both antibody and cellular responses to measles immunization [51]. The links between TLR3 and measles vaccination immunity are particularly intriguing because TLR3 had been found to be a main target for laboratory-adapted measles virus strains, but not for wild-type measles virus strains, in the generation of host immunity. Lab-adapted and vaccine-derived measles virus isolates, such as Edmonston, up-regulate TLR3 expression in human dendritic cells through enhanced IFN-β release according to Tanabe et al. (2003) [50].

Exon 1 is 500 bp upstream the region of the *TLR3* gene that reacts to the measles virus. The region contains the binding sites for the transcription factors NF-kB and STAT (a family of eukaryotic transcription factors that mediates the response to a variety of cytokines and growth factors) and the interferon-stimulated response element (ISRE), also known as an IFN-β induction site, is located 30 bp upstream of exon 1 in the promoter region of the *TLR3* gene [51].

Once the IFN response is activated, the transcript levels of more than 300 genes (ISGs) rise. These genes create proteins with immunoregulatory and antiviral properties that, in some cases, can stop the spread of viruses and reduce their growth. MxA, OAS, and A3G are recognized as reliable markers of IFN activity and are particularly important in the immunological intracellular response. One of the molecular mechanisms by which HTLV-1 counteracts and evades the IFN system is the cellular protein SOCS, which suppresses STAT1 phosphorylation and blocks intracellular signal transduction downstream of the IFN receptor, IFNAR1/2, in CD4+ cells from HTLV-1-infected individuals. When cells infected with HTLV-1 are cultured in vitro, Tyk2 and STAT2, two essential molecules in the initiation chain of the IFN pathway, have lower levels of phosphorylation. Additionally, when p30 is present, the interferon response is suppressed during viral replication, which contributes to the inhibition of TLR3 signaling [52].

These differential outcomes between infections and this SNP in different global populations can be traced to each individual's particular immune response, which is closely related to his or her genetic background. Studies in immunogenetics are currently very helpful to determine the roles in the vulnerability of infection prevention, as they consider genetic background variables [53]. This highlights the need for more research to elucidate the precise function of SNP alleles in the susceptibility or protection of viral diseases.

Regarding SARS-CoV-2, Dhangadamajhi et al. (2021) examined the potential relationship between this *TLR3* variant and COVID-19 based on open international genomic databases of world populations and came to the conclusion that SNP is related to susceptibility to disease and mortality [54]. Although flaws in the analysis were discovered, Pati et al. (2021) revealed these facts in a publication so that the scientific community can carefully assess the findings of the previous article [55]. A five-year-old Brazilian male patient who was the subject of another case study was thought to be susceptible to hepatitis C due to SARS-CoV-2, but the genotyping of the SNP under discussion in this analysis failed to identify the mutant allele (T) in this person [56]. Due to the applied methodology, letters to the editor and case studies were excluded from this meta-analysis.

This present meta-analysis is a pioneering investigation of the evaluation of infections, in general. However, previous meta-analyses have already examined associations between this SNP and particular infections, including one that reported a relationship between this SNP and HIV-1 (which played a protective role for this infection) [57]; the significant association between the mutant allele and the risk of HBV infection [58]; an estimate of an almost two-fold increase in the risk effect for both HBV and HCV infections [59].

This study is constrained by the following factors: (1) the variation in definitions of disease cases from various studies; (2) the heterogeneity of the SNP acting as a potential bias in characteristics such as ethnicities and ages of different populations due to the phenomenon of a genetic background; (3) the methodology used; (4) the need to evaluate the data of interactions between gene and environmental factors.

5. Conclusions

The SNP of the *TLR3* gene, rs3775291, which is related to viral infections, has been linked to the protection, susceptibility, and severity of a number of diseases. Although this polymorphism has been extensively investigated, there is debate over which allele would be linked to the severe form of the illness, as in the case of studies on TBEV infection.

Despite this, the meta-analysis allows the verification of SNP behavior between populations, which can be linked to susceptibility in the cases of Americans and Asians or protection against disease in the case of those living on the European continent. To further understand the role of this SNP in infection and the development of novel pharmaceutical medicines that aid in the treatment. Further research focusing on various diseases, particularly those caused by viruses and the analysis of people from other countries, may be helpful.

Author Contributions: Conceptualization, M.J.A.S. and L.N.G.C.L.; Methodology, M.J.A.S.; software and resources, P.A.S.d.S.; formal analysis, M.J.A.S. and C.S.S.; data curation, M.J.A.S., C.S.S. and M.C.d.S.V.; investigation, M.J.A.S., C.S.S. and M.C.d.S.V.; validation, M.J.A.S., P.A.S.d.S. and C.C.F.; visualization, M.J.A.S., K.V.B.L. and L.N.G.C.L.; roles/writing—original draft and editing, M.J.A.S. and C.S.S.; writing—review and editing, C.C.F., K.V.B.L. and L.N.G.C.L.; supervision, K.V.B.L. and L.N.G.C.L.; project administration, L.N.G.C.L. All authors have read and agreed to the published version of the manuscript.

Funding: The funders had no role in the study design, data collection and analysis, decision to publish, or manuscript preparation.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions of the study are included in the article. Further inquiries can be directed to the corresponding authors.

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

References

1. Borges, G.M. Health transition in Brazil: Regional variations and divergence/convergence in mortality. *Cad. Saúde Pública* **2017**, *33*, e00080316. [CrossRef] [PubMed]
2. World Health Organization (Ed.) *The World Health Report 2000: Health Systems: Improving Performance*; World Health Organization: Geneva, Switzerland, 2000; ISBN 92-4-156198-X.
3. Gadelha, M.R.; Kasuki, L.; Korbonits, M. The genetic background of acromegaly. *Pituitary* **2017**, *20*, 10–21. [CrossRef] [PubMed]
4. Sánchez, D.; Lefebvre, C.; Rioux, J.; García, L.F.; Barrera, L.F. Evaluation of Toll-like receptor and adaptor molecule polymorphisms for susceptibility to tuberculosis in a Colombian population. *Int. J. Immunogenet.* **2012**, *39*, 216–223. [CrossRef]
5. Rai, R.C. Host inflammatory responses to intracellular invaders: Review study. *Life Sci.* **2020**, *240*, 117084. [CrossRef]
6. Teräsjärvi, J.; Hakanen, A.; Korppi, M.; Nuolivirta, K.; Gröndahl-Yli-Hannuksela, K.; Mertsola, J.; Peltola, V.; He, Q. Rapid detection of functional gene polymorphisms of TLRs and IL-17 using high resolution melting analysis. *Sci. Rep.* **2017**, *7*, 41522. [CrossRef]

7. Alseoudy, M.M.; Elgamal, M.; Abdelghany, D.A.; Borg, A.M.; El-Mesery, A.; Elzeiny, D.; Hammad, M.O. Prognostic impact of toll-like receptors gene polymorphism on outcome of COVID-19 pneumonia: A case-control study. *Clin. Immunol.* **2022**, *235*, 108929. [CrossRef] [PubMed]
8. Moresco, E.M.Y.; LaVine, D.; Beutler, B. Toll-like receptors. *Curr. Biol.* **2011**, *21*, R488–R493. [CrossRef]
9. Ribet, D.; Cossart, P. How bacterial pathogens colonize their hosts and invade deeper tissues. *Microbes Infect.* **2015**, *17*, 173–183. [CrossRef]
10. Castro, F.A.; Försti, A.; Buch, S.; Kalthoff, H.; Krauss, C.; Bauer, M.; Egberts, J.; Schniewind, B.; Broering, D.C.; Schreiber, S.; et al. TLR-3 polymorphism is an independent prognostic marker for stage II colorectal cancer. *Eur. J. Cancer* **2011**, *47*, 1203–1210. [CrossRef]
11. Chen, Y.; Lin, J.; Zhao, Y.; Ma, X.; Yi, H. Toll-like receptor 3 (TLR3) regulation mechanisms and roles in antiviral innate immune responses. *J. Zhejiang Univ. Sci. B* **2021**, *22*, 609–632. [CrossRef]
12. Caetano, A.R. SNP markers: Basic concepts, applications in animal breeding and management and perspectives for the future. *R. Bras. Zootec.* **2009**, *38*, 64–71. [CrossRef]
13. Habibabadi, H.M.; Parsania, M.; Pourfathollah, A.A.; Haghighat, S.; Sharifi, Z. Association of TLR3 single nucleotide polymorphisms with susceptibility to HTLV-1 infection in Iranian asymptomatic blood donors. *Rev. Soc. Bras. Med. Trop.* **2020**, *53*, e20200026. [CrossRef]
14. Redondo, N.; Rodríguez-Goncer, I.; Parra, P.; Ruiz-Merlo, T.; López-Medrano, F.; González, E.; Polanco, N.; Trujillo, H.; Hernández, A.; San Juan, R.; et al. Influence of single-nucleotide polymorphisms in TLR3 (rs3775291) and TLR9 (rs352139) on the risk of CMV infection in kidney transplant recipients. *Front. Immunol.* **2022**, *13*, 929995. [CrossRef] [PubMed]
15. Stroup, D.F. Meta-analysis of Observational Studies in EpidemiologyA Proposal for Reporting. *JAMA* **2000**, *283*, 2008. [CrossRef] [PubMed]
16. Tugwell, P.; Tovey, D. PRISMA 2020. *J. Clin. Epidemiol.* **2021**, *134*, A5–A6. [CrossRef]
17. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, n71. [CrossRef] [PubMed]
18. Tarsilla, M. Cochrane handbook for systematic reviews of interventions. *J. Multidiscip. Eval.* **2010**, *6*, 142–148. [CrossRef]
19. Wang, B.-G.; Yi, D.-H.; Liu, Y.-F. TLR3 gene polymorphisms in cancer: A systematic review and meta-analysis. *Chin. J. Cancer* **2015**, *34*, 19. [CrossRef]
20. Tang, J.-L.; Liu, J.L. Misleading funnel plot for detection of bias in meta-analysis. *J. Clin. Epidemiol.* **2000**, *53*, 477–484. [CrossRef]
21. Khanmohammadi, S.; Rezaei, N. Role of Toll-like receptors in the pathogenesis of COVID-19. *J. Med. Virol.* **2021**, *93*, 2735–2739. [CrossRef]
22. Kawai, T.; Akira, S. The role of pattern-recognition receptors in innate immunity: Update on Toll-like receptors. *Nat. Immunol.* **2010**, *11*, 373–384. [CrossRef] [PubMed]
23. Lester, S.N.; Li, K. Toll-Like Receptors in Antiviral Innate Immunity. *J. Mol. Biol.* **2014**, *426*, 1246–1264. [CrossRef] [PubMed]
24. Ishizaki, Y.; Takemoto, M.; Kira, R.; Kusuhara, K.; Torisu, H.; Sakai, Y.; Sanefuji, M.; Yukaya, N.; Hara, T. Association of toll-like receptor 3 gene polymorphism with subacute sclerosing panencephalitis. *J. Neurovirol.* **2008**, *14*, 486–491. [CrossRef] [PubMed]
25. Ranjith-Kumar, C.T.; Miller, W.; Sun, J.; Xiong, J.; Santos, J.; Yarbrough, I.; Lamb, R.J.; Mills, J.; Duffy, K.E.; Hoose, S.; et al. Effects of single nucleotide polymorphisms on Toll-like receptor 3 activity and expression in cultured cells. *J. Biol. Chem.* **2007**, *282*, 17696–17705. [CrossRef] [PubMed]
26. Gao, D.; Wang, X.D.; Jin, G.Y. Research progress of Toll-like receptor and ligand complex structure. *China J. Cell. Mol. Immunol.* **2015**, *31*, 553–556.
27. Qi, R.; Hoose, S.; Schreiter, J.; Sawant, K.V.; Lamb, R.; Ranjith-Kumar, C.T.; Mills, J.; San Mateo, L.; Jordan, J.L.; Kao, C.C. Secretion of the human Toll-like receptor 3 ectodomain is affected by single nucleotide polymorphisms and regulated by Unc93b1. *J. Biol. Chem.* **2010**, *285*, 36635–36644. [CrossRef]
28. Yang, Q.; Shu, H.-B. Deciphering the pathways to antiviral innate immunity and inflammation. *Adv. Immunol.* **2020**, *145*, 1–36.
29. Matsumoto, M.; Oshiumi, H.; Seya, T. Antiviral responses induced by the TLR3 pathway. *Rev. Med. Virol.* **2011**, *21*, 67–77. [CrossRef]
30. Redondo, N.; Rodríguez-Goncer, I.; Parra, P.; López-Medrano, F.; González, E.; Hernández, A.; Trujillo, H.; Ruiz-Merlo, T.; San Juan, R.; Folgueira, M.D.; et al. Genetic polymorphisms in TLR3, IL10 and CD209 influence the risk of BK polyomavirus infection after kidney transplantation. *Sci. Rep.* **2022**, *12*, 11338. [CrossRef]
31. Lee, S.-O.; Brown, R.A.; Razonable, R.R. Association between a functional polymorphism in Toll-like receptor 3 and chronic hepatitis C in liver transplant recipients. *Transpl. Infect. Dis.* **2013**, *15*, 111–119. [CrossRef]
32. Santos, C.N.O.; Ribeiro, D.R.; Cardoso Alves, J.; Cazzaniga, R.A.; Magalhães, L.S.; de Souza, M.S.F.; Fonseca, A.B.L.; Bispo, A.J.B.; Porto, R.L.S.; dos Santos, C.A.; et al. Association Between Zika Virus Microcephaly in Newborns With the rs3775291 Variant in Toll-Like Receptor 3 and rs1799964 Variant at Tumor Necrosis Factor-α Gene. *J. Infect. Dis.* **2019**, *220*, 1797–1801. [CrossRef] [PubMed]
33. Posadas-Mondragón, A.; Aguilar-Faisal, J.L.; Zuñiga, G.; Magaña, J.J.; Santiago-Cruz, J.A.; Guillén-Salomón, E.; Alcántara-Farfán, V.; Arellano-Flores, M.L.; Salas-Benito, J.S.; Neri-Bazán, R.M.; et al. Association of Genetic Polymorphisms in TLR3, TLR4, TLR7, and TLR8 with the Clinical Forms of Dengue in Patients from Veracruz, Mexico. *Viruses* **2020**, *12*, 1230. [CrossRef] [PubMed]

34. Li, G.; Zheng, Z. Toll-like receptor 3 genetic variants and susceptibility to hepatocellular carcinoma and HBV-related hepatocellular carcinoma. *Tumour Biol.* **2013**, *34*, 1589–1594. [CrossRef] [PubMed]
35. Rong, Y.; Song, H.; You, S.; Zhu, B.; Zang, H.; Zhao, Y.; Li, Y.; Wan, Z.; Liu, H.; Zhang, A.; et al. Association of Toll-like receptor 3 polymorphisms with chronic hepatitis B and hepatitis B-related acute-on-chronic liver failure. *Inflammation* **2013**, *36*, 413–418. [CrossRef] [PubMed]
36. Alagarasu, K.; Bachal, R.V.; Memane, R.S.; Shah, P.S.; Cecilia, D. Polymorphisms in RNA sensing toll like receptor genes and its association with clinical outcomes of dengue virus infection. *Immunobiology* **2015**, *220*, 164–168. [CrossRef]
37. Biyani, S.; Garg, R.K.; Jain, A.; Malhotra, H.S.; Kumar, R.; Prakash, S.; Verma, R.; Sharma, P.K. Toll-like receptor-3 gene polymorphism in patients with Japanese encephalitis. *J. Neuroimmunol.* **2015**, *286*, 71–76. [CrossRef]
38. Wan, Z.; Lin, X.; Li, T.; Zhou, A.; Yang, M.; Hu, D.; Feng, L.; Peng, S.; Fan, L.; Tu, S.; et al. Genetic variant in CXCL13 gene is associated with susceptibility to intrauterine infection of hepatitis B virus. *Sci. Rep.* **2016**, *6*, 26465. [CrossRef]
39. Chen, D.; Xie, W.; Lu, Y.; Su, S.; Nong, L.; Jia, Y.; Liu, Y.; Zhou, W.; Wang, H.; Tan, A. Gene polymorphisms of TLR2 and TLR3 in HBV clearance and HBV-related hepatocellular carcinoma in a Chinese male population. *Int. J. Biol. Mrk.* **2017**, *32*, e195–e201. [CrossRef]
40. Singh, A.K.; Prakash, S.; Garg, R.K.; Jain, P.; Kumar, R.; Jain, A. Study of Single Nucleotide Polymorphisms in Endosomal Toll-Like Receptors-3, 7, and 9 Genes in Patients With Dengue: A Case-Control Study. *Cureus* **2021**, *13*, e14883. [CrossRef]
41. Préhaud, C.; Mégret, F.; Lafage, M.; Lafon, M. Virus Infection Switches TLR-3-Positive Human Neurons To Become Strong Producers of Beta Interferon. *J. Virol.* **2005**, *79*, 12893–12904. [CrossRef]
42. Huik, K.; Avi, R.; Pauskar, M.; Kallas, E.; Jõgeda, E.-L.; Karki, T.; Marsh, K.; Des Jarlais, D.; Uuskūla, A.; Lutsar, I. Association between TLR3 rs3775291 and resistance to HIV among highly exposed Caucasian intravenous drug users. Infect. Genet. Evol. J. Mol. Epidemiol. *Evol. Genet. Infect. Dis.* **2013**, *20*, 78–82. [CrossRef] [PubMed]
43. Mickienė, A.; Pakalnienė, J.; Nordgren, J.; Carlsson, B.; Hagbom, M.; Svensson, L.; Lindquist, L. Polymorphisms in chemokine receptor 5 and Toll-like receptor 3 genes are risk factors for clinical tick-borne encephalitis in the Lithuanian population. *PLoS ONE* **2014**, *9*, e106798. [CrossRef]
44. Svensson, A.; Tunbäck, P.; Nordström, I.; Padyukov, L.; Eriksson, K. Polymorphisms in Toll-like receptor 3 confer natural resistance to human herpes simplex virus type 2 infection. *J. Gen. Virol.* **2012**, *93*, 1717–1724. [CrossRef] [PubMed]
45. Reinert, L.S.; Harder, L.; Holm, C.K.; Iversen, M.B.; Horan, K.A.; Dagnæs-Hansen, F.; Ulhøi, B.P.; Holm, T.H.; Mogensen, T.H.; Owens, T.; et al. TLR3 deficiency renders astrocytes permissive to herpes simplex virus infection and facilitates establishment of CNS infection in mice. *J. Clin. Investig.* **2012**, *122*, 1368–1376. [CrossRef] [PubMed]
46. Guo, Y.; Audry, M.; Ciancanelli, M.; Alsina, L.; Azevedo, J.; Herman, M.; Anguiano, E.; Sancho-Shimizu, V.; Lorenzo, L.; Pauwels, E.; et al. Herpes simplex virus encephalitis in a patient with complete TLR3 deficiency: TLR3 is otherwise redundant in protective immunity. *J. Exp. Med.* **2011**, *208*, 2083–2098. [CrossRef]
47. Wu, J.; Lu, M.; Meng, Z.; Trippler, M.; Broering, R.; Szczeponek, A.; Krux, F.; Dittmer, U.; Roggendorf, M.; Gerken, G.; et al. Toll-like receptor-mediated control of HBV replication by nonparenchymal liver cells in mice. *Hepatology* **2007**, *46*, 1769–1778. [CrossRef]
48. Horscroft, N.J.; Pryde, D.C.; Bright, H. Antiviral applications of Toll-like receptor agonists. *J. Antimicrob. Chemother.* **2012**, *67*, 789–801. [CrossRef]
49. Levi, J.E. Emerging Infectious Agents and Blood Safety in Latin America. *Front. Med.* **2018**, *5*, 71. [CrossRef]
50. Tanabe, M.; Kurita-Taniguchi, M.; Takeuchi, K.; Takeda, M.; Ayata, M.; Ogura, H.; Matsumoto, M.; Seya, T. Mechanism of up-regulation of human Toll-like receptor 3 secondary to infection of measles virus-attenuated strains. *Biochem. Biophys. Res. Commun.* **2003**, *311*, 39–48. [CrossRef]
51. Dhiman, N.; Ovsyannikova, I.G.; Vierkant, R.A.; Ryan, J.E.; Pankratz, V.S.; Jacobson, R.M.; Poland, G.A. Associations Between SNPs in Toll-like Receptors and Related Intracellular Signaling Molecules and Immune Responses to Measles Vaccine: Preliminary Results. *Vaccine* **2008**, *26*, 1731–1736. [CrossRef]
52. Fenizia, C.; Fiocchi, M.; Jones, K.; Parks, R.W.; Ceribelli, M.; Chevalier, S.A.; Edwards, D.; Ruscetti, F.; Pise-Masison, C.A.; Franchini, G. Human T-Cell Leukemia/Lymphoma Virus Type 1 p30, but Not p12/p8, Counteracts Toll-Like Receptor 3 (TLR3) and TLR4 Signaling in Human Monocytes and Dendritic Cells. *J. Virol.* **2014**, *88*, 393–402. [CrossRef] [PubMed]
53. Doetschman, T. Influence of Genetic Background on Genetically Engineered Mouse Phenotypes. *Methods Mol. Biol.* **2009**, *530*, 423–433. [CrossRef] [PubMed]
54. Dhangadamajhi, G.; Rout, R. Association of TLR3 functional variant (rs3775291) with COVID-19 susceptibility and death: A population-scale study. Hum. *Cell* **2021**, *34*, 1025–1027. [CrossRef] [PubMed]
55. Pati, A.; Padhi, S.; Chaudhury, S.; Panda, A.K. TLR3 (rs3775291) variant is not associated with SARS-CoV-2 infection and related mortality: A population-based correlation analysis. Hum. *Cell* **2021**, *34*, 1274–1277. [CrossRef]
56. Pessoa, N.L.; Bentes, A.A.; de Carvalho, A.L.; de Souza Silva, T.B.; Alves, P.A.; de Sousa Reis, E.V.; Rodrigues, T.A.; Kroon, E.G.; Campos, M.A. Case report: Hepatitis in a child infected with SARS-CoV-2 presenting toll-like receptor 7 Gln11Leu single nucleotide polymorphism. *Virol. J.* **2021**, *18*, 1–5. [CrossRef]
57. Shi, H.; He, H.; Sun, C.; Fu, J.; Ghosh, D.; Deng, C.; Sheng, Y. Association of toll-like receptor polymorphisms with acquisition of HIV infection and clinical findings. *Medicine* **2020**, *99*, e23663. [CrossRef]

58. Ye, S.; Zhang, X.; Zhang, Y.; Tian, X.; Liu, A.; Cui, C.; Shi, L.; Xia, D. Association of TLR3 (rs3775291) and IL-10 (rs1800871) gene polymorphisms with susceptibility to Hepatitis B infection: A meta-analysis. *Epidemiol. Infect.* **2020**, *148*, e228. [CrossRef]
59. Geng, P.-L.; Song, L.-X.; An, H.; Huang, J.-Y.; Li, S.; Zeng, X.-T. Toll-Like Receptor 3 is Associated With the Risk of HCV Infection and HBV-Related Diseases. *Medicine* **2016**, *95*, e2302. [CrossRef]

APENDICE 2

genes

MDPI

Article

Analysis of Epidemiological Factors and SNP rs3804100 of *TLR2* for COVID-19 in a Cohort of Professionals Who Worked in the First Pandemic Wave in Belém-PA, Brazil

Marcos Jessé Abrahão Silva [1,*], Caroliny Soares Silva [2], Rebecca Lobato Marinho [3], Jeanne Gonçalves Cabral [3], Ellen Polyana da Costa Gurrão [3], Pabllo Antonny Silva dos Santos [2], Samir Mansour Moraes Casseb [3], Karla Valéria Batista Lima [3] and Luana Nepomuceno Gondim Costa Lima [3,*]

[1] Master Program in Epidemiology and Health Surveillance (PPGEVS), Evandro Chagas Institute (IEC), Ananindeua 67030-000, PA, Brazil
[2] Master and PhD Program in Parasitic Biology in the Amazon (PPGBPA), University of State of Pará (UEPA), Belém 66087-670, PA, Brazil; karolinysoares2303@gmail.com (C.S.S.); antonnypabllo@gmail.com (P.A.S.d.S.)
[3] Evandro Chagas Institute (IEC), Ananindeua 67030-000, PA, Brazil; rebeccamarinho28@gmail.com (R.L.M.); jeannegoncalves@iec.gov.br (J.G.C.); ellengurrao@iec.gov.br (E.P.d.C.G.); samircasseb@ufpa.br (S.M.M.C.); karlalima@iec.gov.br (K.V.B.L.)
* Correspondence: jesseabrahao10@gmail.com (M.J.A.S.); luanalima@iec.gov.br (L.N.G.C.L.)

Abstract: COVID-19 is an infectious disease caused by coronavirus 2 of the severe acute syndrome (SARS-CoV-2). Single nucleotide polymorphisms (SNPs) in genes, such as *TLR2*, responsible for an effective human immune response, can change the course of infection. The objective of this article was to verify associations between epidemiological factors and *TLR2* SNP rs3804100 (Thymine [T] > Cytosine [C]) in professionals from Health Institutions (HI) who worked during the first pandemic wave and COVID-19. A case-control study was conducted with Belém-PA HI workers (Northern Brazil), divided into symptomatology groups (Asymptomatic-AS; $n = 91$; and Symptomatic-SI; $n = 123$); and severity groups classified by Chest Computerized Tomography data (symptomatic with pulmonary involvement—SCP; $n = 35$; symptomatic without pulmonary involvement—SSP; $n = 8$). Genotyping was performed by Sanger sequencing, and Statistical Analysis was conducted through the SPSS program. Bioinformatics servers predicted the biological functions of the *TLR2* SNP. There were associations between the presence of comorbidities and poor prognosis of COVID-19 (especially between symptomatology and severity of COVID-19 and overweight and obesity) and between the sickness in family members and kinship (related to blood relatives). The homozygous recessive (C/C) genotype was not found, and the frequency of the mutant allele (C) was less than 10% in the cohort. No significant associations were found for this SNP in this cohort. The presence of SNP was indicated to be benign and causes a decrease in the stability of the TLR2 protein. These data can help the scientific community and medicine find new forms of COVID-19 containment.

Keywords: *TLR2*; single nucleotide polymorphism; COVID-19; epidemiology

check for updates

Citation: Silva, M.J.A.; Silva, C.S.; Marinho, R.L.; Cabral, J.G.; Gurrão, E.P.d.C.; dos Santos, P.A.S.; Casseb, S.M.M.; Lima, K.V.B.; Lima, L.N.G.C. Analysis of Epidemiological Factors and SNP rs3804100 of *TLR2* for COVID-19 in a Cohort of Professionals Who Worked in the First Pandemic Wave in Belém-PA, Brazil. *Genes* **2023**, *14*, 1907. https://doi.org/10.3390/genes14101907

Academic Editor: Zsolt Ronai

Received: 17 August 2023
Revised: 19 September 2023
Accepted: 25 September 2023
Published: 5 October 2023

1. Introduction

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a species of beta-coronavirus. Zoonotic-originating disease (probably due to bat overflow) in most cases typically causes signs and symptoms similar to those of influenza, and in about 10 to 20% of cases, it can lead to pneumonia through acute respiratory distress syndrome (ARDS) for 8 to 14 days with dyspnea and reaching hypoxemia, in addition to being asymptomatic [1].

Due to its high transmissibility (through expelled respiratory droplets, aerosols, surface contamination, and fomites), the disease spread rapidly throughout the world [2]. In this context, the World Health Organization (WHO) classified it as a pandemic and public

health emergency of international concern just three months after the identification of the first case [3]. After the notification of the pandemic, more than 500 million confirmed cases and about 6 million deaths were registered worldwide, with two waves that had already been notified in several countries, such as Brazil, in some cases, a third wave and a possible fourth wave were announced due to the new variants of the virus [4]. Thus, the number of cases of COVID-19 continues to rise.

Brazil is the second country with the highest number of deaths from the disease [4]. According to epidemiologists, the first wave of COVID-19 in Brazil began on 25 February 2020 and ended around 31 October 2020 [5,6]. In particular, in Brazil, the first wave of the pandemic was characterized by the difficulty in detecting asymptomatic individuals and the deficient policy of availability of diagnostic tests in the country with respect to mass tests for COVID-19, due to the lack of central political leadership in the provision of health networks in the planning and promotion of their management [7,8].

Reference diagnosis is based on reverse transcription polymerase chain reaction (RT-PCR), and the association between laboratory and tomographic findings by Chest Computed Tomography (CCT) may be present, even in initial symptomatic patients [9–11]. In the first wave of the pandemic, due to the fragile system of availability of confirmatory tests, the considerable increase in cases, and the proximity of disease symptoms to flu syndromes, the WHO recommended that people who had at least two of the suggestive symptoms should already be classified as cases of the disease, which was also adopted by the Brazilian Ministry of Health [12,13].

Different clinical manifestations and evolutions of COVID-19 can be related to several factors, from the viral amount to which the individual is exposed (viral load) to sociodemographic, behavioral, immunological, and genetic aspects of the host [14]. The immune system plays a key role in the fight against Severe Acute Respiratory Syndrome—SARS [15]. In fact, the innate immune system is capable of recognizing the molecular structures produced by SARS-CoV-2 infection (by virus invasion) [16].

From this perspective, *Toll-like Receptors* (*TLRs*) are genes discovered, initially in the 1980s in flies of the genus Drosophila and later in mammals, in which they encode transmembrane proteins, characterized by the presence of an extracellular domain N-terminal rich in leucine repeats (LRR) and acting as pattern recognition receptors (PRRs) for the development and activation of the innate immune system by recognizing pathogen-associated molecular patterns (PAMPs) of invading agents. Each TLR can recognize specific types of PAMPs and/or Damage-Associated Molecular Patterns (DAMPs) [17].

In the case of SARS-CoV-2 infection, TLR2 can track beta (β)-coronavirus infection through recognition of the E protein, inducing the release of pro-inflammatory cytokines such as TNF-α and interferon-gamma (IFN-γ) [18]. Therefore, there is significant cytokine production in the body and a contribution to the generation of adaptive immunity by monitoring the expression of costimulatory molecules for defense against pathogens, as in the case of SARS-CoV-2 [19].

The immunogenetic aspects of the host in relation to infection can be investigated, for example, through single nucleotide polymorphisms (SNPs), which encode an amino acid alteration in genes responsible for an effective human immune response, thus altering the course of viral infection [20]. It is necessary to understand the importance of host immunogenetic heterogeneity in the *TLR2* gene for the evolution of COVID-19 to create strategies to cope with the disease. *TLRs* play a key role in generating and maintaining the innate immune response, in addition to guiding the adaptive immune response in infections. Therefore, SNPs in these genes can generate different levels of receptor expression, generating uncoordinated immune responses through excessive or reduced cytokine production [21–23].

Tuberculosis (TB), an essentially lung disease, was the disease most studied for the SNP rs3804100 in the *TLR2* gene in different populations [21]. This may be linked to the fact that the highlighted gene is the main receptor for lipoproteins in mammals, derived from a variety of bacteria, such as the agent *Mycobacterium tuberculosis*. In addition, TLR2 is a

crucial factor in activating IFN-γ, which promotes cellular immunity in the Th1 population through the induction of IL-12 [24–26]. The main point of similarity between tuberculosis and COVID-19 is the fact that both affect the respiratory system of the infected person. In this context, it is scientifically valid to evaluate the same SNP and its associations for COVID-19.

Therefore, in this study, we propose an evaluation of epidemiological factors and SNP rs3804100 (Thymine to Cytosine, T to C) in the *TLR2* gene of Health Institutions professionals in the city of Belém (capital of the State of Pará, Brazil) who worked during the first pandemic wave to analyze the relationship with symptomatology and clinical development of COVID-19. The use of the first wave of COVID-19 in this work is due to the fact that there were still no variants of the virus that could resort to biases in the analysis of the evolution of the condition of these individuals, as well as the fact that the genetic data analyzed of these individuals reflect the observation that genetic characteristics, such as an SNP of the *TLR2* gene, do not change in an individual during his or her course of life [27].

2. Material and Methods

2.1. Study Design and Ethical Considerations

This study is characterized as observational with a quantitative analytical nature and is classified as a retrospective case-control. It followed the recommendations of Strengthening the reporting of observational studies in epidemiology (STROBE) [28]. This study protocol was approved by local ethics committees, and all subjects gave their written informed consent (Term of Free and Informed Consent—TCLE). This work was approved by the Research Ethics Committee of the State University of Pará—UEPA (CAAE: 38113620.1.0000.5174) and is related to the research project "Análise da resposta ao SARS-CoV-2 em relação aos achados radiológicos e/ou à susceptibilidade genética individual", with opinion number: 6.124.862. This research was carried out in accordance with the Helsinki Declaration [29] and Resolution Nº. 466/2012 of the Brazilian National Health Council [30].

2.2. Settings and Participants

This study was carried out in 10 health institutions located in the city of Belém-PA, Brazil (in the Amazon Region—Northern Brazil). They were: Policlínica Metropolitana de Belém (PMB); Jean Bittar Hospital; Hospital das Clínicas Gaspar Vianna; João de Barros Barreto University Hospital; Hospital Adventista de Belém; Dom Vicente Zico Hospital; Women's Health Hospital; Institute of Hematology and Hemotherapy of Belém (IHEBE); Psychosocial Care Center (CAPS); Secretariat of Public Health (SESPA). Institutions were chosen randomly based on obtaining contacts that fit this study population. All participating institutions received subjects infected with COVID-19 during the period of the first wave of the pandemic.

This study included, through a convenience sampling, 214 health, administration, and general services professionals who actively worked in health institutions that received individuals with COVID-19 in the period between 1 April 2021 and 30 June 2020, exposed directly and daily to SARS-CoV-2, since during this period all individuals who worked in health institutions can be considered directly exposed to SARS-CoV-2 [31]. During this period, all people who worked in health institutions can be considered exposed to SARS-CoV-2 since safety protocols were not yet well established, there was overcrowding in health institutions, masks were scarce, and the N95 mask was not used [32].

2.3. Variables and Division of Cohort in Groups of This Study

These professionals were first divided into two groups. Group 1, which is made up of individuals who were in constant contact with a patient with COVID-19 throughout the first wave but did not report symptoms of COVID-19, was called asymptomatic (AS), and group 2, composed of professionals who were also in direct contact and reported at least two characteristic symptoms of COVID-19 (with or without COVID-19 testing), was designated as symptomatic (SI).

The presentation of two of the main symptoms suggestive of COVID-19 infection and that these symptoms were related to the characteristic clinical presentation of the first wave of COVID-19 (dyspnea, fever, or dry cough) was considered a symptomatological illness criterion [33,34].

The second division, from the perspective of aggravation, is composed of groups 3 and 4, which are individuals from group 2, who performed Chest Computerized Tomography (CCT). Group 3 consisted of individuals who reported at least two symptoms and also reported pulmonary impairment ≥ 10%, named symptomatic with pulmonary involvement (called SCP), and group 4 also reported at least two symptoms of COVID-19 and had a CCT with a result without lung injury, this group was named symptomatic without pulmonary involvement (called SSP). The evaluation of severity by CCT in COVID-19 is what was recommended by the Ministry of Health and WHO [35,36].

Questionnaires developed by this study group were evaluated for epidemiological analysis and possible correlation with the genetic data of these individuals. Information such as signs and symptoms related to the period of the first wave, sex, age, occupation, and comorbidities was used. CCT data were added based on the patient's self-report while completing the questionnaire. These data were collected from 1 November 2021 to 1 November 2022. Professionals in health, administrative, and general service areas who did not work from April to June 2020 in places that directly assisted patients with COVID-19 or individuals who did not accept to participate in the research or who did not sign the consent form were excluded from the survey. The sampling and its appropriate groups are represented in Figure 1.

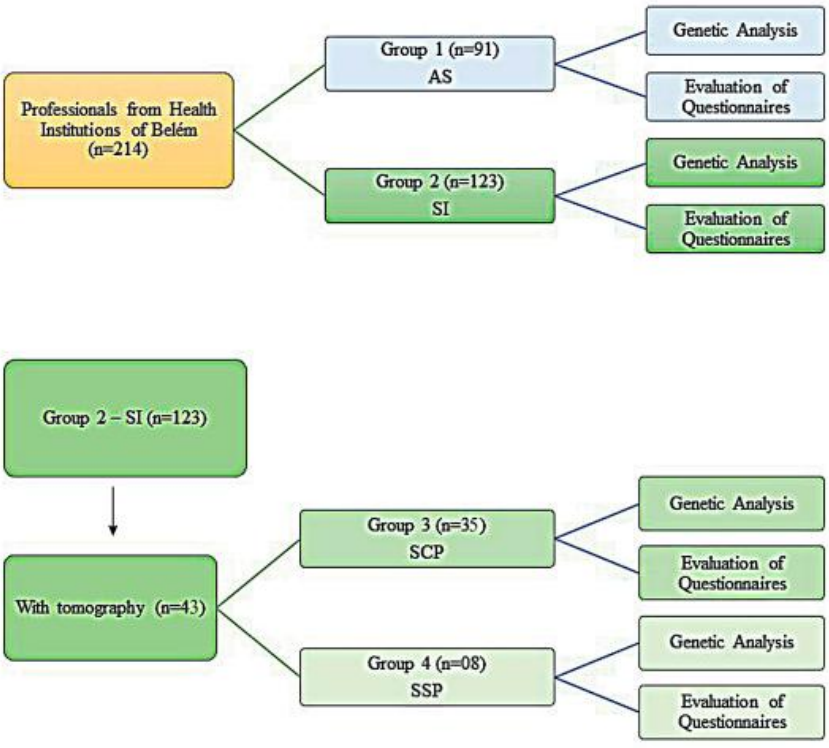


Figure 1. Sample Processing and Classification Flowchart.

2.4. Sample Collection, DNA Isolation, and Amplification of the Samples by Polymerase Chain Reaction (PCR)

Sample collection was carried out between 1 June 2021 and 30 March 2022. Blood samples were taken from professionals who got sick and who did not get sick by venipuncture in 5 mL EDTA tubes, stored at −20 °C for later laboratory procedures in the Molecular Biology Laboratory—LABIMOL, Bacteriology and Mycology Section (SABMI) of the Evandro Chagas Institute (IEC). DNA extractions were performed using the DNeasy Blood and Tissue Kit (QIAGEN®, Venlo, The Netherlands), following the manufacturer's instructions. For all these professionals, the SNP rs3804100 of *TLR2* was evaluated to correlate with the susceptibility and severity of COVID-19. The SNP information (SNP ID) of *TLR2* was retrieved from the National Center for Biotechnology Information—NCBI, dbSNP (http://www.ncbi.nlm.nih.gov/snp/; accessed on 1 January 2022) [37].

The typification of the SNP of the *TLR2* gene was carried out by sequencing through amplification of the DNA by the nucleotide primers (primers) for the Polymerase Chain Reaction—PCR, which were designed by the Primer3Plus program version 2.0 (http://www.bioinformatics.nl/primer3plus/; accessed on 10 December 2021) [38] from the respective genomic regions deposited in GenBank [39]. Thus, 1 primer (with Forward Strand—ACCGGAGAGACTTTGCTCAC and Reverse Strand—GCTTGCTGCTCCTGAGTGAA, 437 base pairs) at the 5′ Binding site 153704120 and reference NC_000004.12 were designed for use and amplification [37]. They were performed with Platinum Taq DNA Polymerase, DNA-free (Invitrogen®, Thermo Fisher Scientific Corporation, Waltham, MA, USA) using the following conditions in a thermocycler: initial denaturation at 95 °C for 1 min, followed by 35 cycles of denaturation at 95 °C to 30 s, annealing at the appropriate temperature for this primer (65 °C) for 30 s, extension at 72 °C to 1 min, and after that, final extension at 72 °C for 10 min [40]. The amplified products were subjected to electrophoresis in a 2% agarose gel containing 3.0 µL of Sybr Safe (Invitrogen®, Thermo Fisher Scientific Corporation, Waltham, MA, USA) to visualize the amplified DNA fragments in a photodocumenting device.

2.5. Running Samples in Capillary Electrophoresis

Purification of PCR products was performed using the EasyPure PCR Purification Kit (TransGen Biotech Co.®, Beijing, China), according to the manufacturer's recommendations. The already purified amplified products were submitted to the sequencing reaction following the instructions of the BigDye X-Terminator kit to the ABI 3130 Genetic Analyzer sequencer (Applied Biosystems®, Life Technologies, Thermo Fisher Scientific Corporation, Waltham, MA, USA) for visualization and analysis of the areas of interest of SNP through the Bioedit program version 7.2.5 [41], respectively, with subsequent performance of BLAST on the NCBI website (https://blast.ncbi.nlm.nih.gov/Blast.cgi/; accessed on 15 August 2023).

2.6. Presentation of Data and Statistical Analysis of Results

Information regarding laboratory results was organized in a database using the Microsoft Office Access program (Microsoft Corp®, Redmond, WA, USA) and presented through graphs or tables generated by the Microsoft Office Excel program (Microsoft Corp.®, Redmond, WA, USA). The observed proportions of the presence of the SNP within each studied group were analyzed with the aid of IBM SPSS Statistics v. 26.0 software (IBM Corp.®, Armonk, NY, USA), using the G, two-tailed chi-square (χ^2), and Fisher's exact tests to verify the association between variables arranged in 2 × 2 tables. An Odds Ratio (OR) test with a 95% confidence interval (CI) was used to assess the association between exposure and the outcomes of interest. A probability (p) ≤ 0.05 was considered statistically significant.

Genotype frequencies were tested for the Hardy-Weinberg Equilibrium (HWE) using the chi-square (χ^2) test with $p < 0.001$ as the cut-off point for the significance level through Arlequin version 3.5.1.2 [42,43]. The G*Power software version 3.1.9.7 was used to

determine the power of the sample size using a goodness-of-fit Chi-squared test [44]. In addition to that, the biological functions of the SNP were evaluated by PolyPhen-2 (Polymorphism Phenotyping, http://genetics.bwh.harvard.edu/pph2/; accessed on 22 January 2023) through the UniProt Database Entry O60603 (Human TLR2) [45], while protein structural stability was evaluated using I-Mutant 2.0 with Protein Data Bank—PDB Code 2Z7X (https://folding.biofold.org/i-mutant/i-mutant2.0.html/; accessed on 27 March 2023) [46].

3. Results

3.1. Power of Sample Size, Normality of Variables, and Hardy-Weinberg Equilibrium (HWE)

The power of the sample size was estimated using Chi-square quality-of-fit for symptomatic (N = 123) and asymptomatic individuals with COVID-19 (N = 91), with an α error probability of 0.05 and an effect size of 0.3. The actual real power (1-β error probability) was 0.99, which is greater than 0.80, i.e., statistically acceptable [44]. The variables analyzed in this study were disposed of in a categorical and nonparametric way, characterized by an absolute count and percentage. The SNP rs3804100 was in agreement with the HWE ($p = 0.93$).

3.2. Baseline Characteristics Associated with COVID-19 Symptomatology among Individuals in the Belém Professional Cohort

The epidemiological characteristics and comorbidities associated with the cohort are presented in Table 1. There were no statistically significant differences between the age group and sex categories and COVID-19 ($p > 0.05$). However, with respect to pre-existing comorbidities in these subjects, the absence of comorbidities was associated with the group of asymptomatic individuals, in which the absence of comorbidity is associated with a two times greater chance of being asymptomatic ($p = 0.0034$; OR = 2.61 [95% CI = 1.35–5.01]). The number of comorbidities was also a significant factor for symptomatology in the individuals, and, therefore, the presentation of two or more comorbidities was statistically significant and was present only in the symptomatic group ($p = 0.020$; OR = 0.66 [95% CI = 0.54–0.81]). The types of comorbidities are shown in Table 1. Among these, the presence of Diabetes Mellitus or Overweight and obesity were associated with the group of people with symptoms of COVID-19 ($p = 0.021$, OR= 0.56 [95% CI = 0.49–0.63]; $p = 0.003$, OR = 4.17 [95% CI = 1.52–11.40], respectively).

Table 1. Epidemiological characteristics of professionals exposed to SARS-CoV-2.

Variable *n* (%)	AS (*n* = 91) *n* (%)	SI (*n* = 123) *n* (%)	*p*-Value
Age group			
19–34 years	40 (44%)	63 (51.2%)	$p > 0.05$
35–50 years	43 (47.2%)	49 (39.9%)	
>50 years	8 (8.8%)	11 (8.9%)	
Sex			
Female	67 (73.6%)	81 (65.9%)	$p > 0.05$
Male	24 (26.4%)	42 (34.1%)	
Presence of pre-existing comorbidities			
No comorbidities	75 (82.4%)	79 (64.2%)	**$p = 0.0034$** [a]
With comorbidities	16 (17.6%)	44 (35.8%)	
Quantity of comorbidities			
1 comorbidity	16 (100%)	32 (72.73%)	**$p = 0.020$** [b]
≥2 comorbidities	0 (0%)	12 (27.27%)	

Table 1. *Cont.*

Variable *n* (%)	AS (*n* = 91) *n* (%)	SI (*n* = 123) *n* (%)	*p*-Value
Types of comorbidities			
Asthma	4 (25%)	11 (18.33%)	$p > 0.05$
Cardiopathies	0 (0%)	4 (6.67%)	$p > 0.05$
Diabetes mellitus	0 (0%)	7 (11.67%)	**$p = 0.021$** [c]
Systemic Arterial Hypertension (SAH)	5 (31.25%)	10 (16.67%)	$p > 0.05$
Overweight and obesity	5 (31.25%)	24 (40%)	**$p = 0.003$** [d]
Autoimmune disease	2 (12.5%)	0 (0%)	$p > 0.05$
Kidney disease	0 (0%)	1 (1.66%)	$p > 0.05$
Pulmonary fibrosis	0 (0%)	1 (1.66%)	$p > 0.05$
Glaucoma	0 (0%)	2 (3.33%)	$p > 0.05$
Profession category			
Administrative	23 (25.3%)	37 (30.1%)	$p > 0.05$
Healthcare professional	56 (61.5%)	65 (52.8%)	
General Services	12 (13.2%)	21 (17.1%)	

Caption: AS = Asymptomatic Group; SI = Symptomatic Group. Note: [a] OR = 2.61 (95% CI = 1.35–5.01); [b] OR = 0.66 (95% CI = 0.54–0.81); [c] OR = 0.56 (95% CI = 0.49–0.63); [d] OR = 4.17 (95% CI = 1.52–11.40).

Table 2 shows the analysis between the AS and SI groups and the illnesses of their family members. Thus, it was found that blood relatives who live with professionals had 2.33 higher chances of becoming ill ($p = 0.016$; OR = 2.33 [95% CI = 1.16–4.69]).

Table 2. Kinship and housing relationships of family members of individuals exposed to SARS-CoV-2.

Variable *n* (%)	Kinship	AS (*n* = 91) *n* (%)	SI (*n* = 123) *n* (%)	*p*-Value
Getting sick				
Relatives who did not get sick	Blood Relatives	38 (41.76%)	40 (32.52%)	$p > 0.05$
Relatives who became ill		53 (58.24%)	83 (67.48%)	
Living with the research participant		24 (45.28%)	58 (69.88%)	**$p = 0.016$** [a]
Not living with the research participant		29 (54.72%)	25 (30.12%)	
Getting sick in the same household				
Relatives who did not get sick	Non-blood relatives	64 (70.3%)	73 (59.3%)	$p > 0.05$
Relatives who became ill		27 (29.7%)	50 (40.7%)	

Caption: AS = Asymptomatic Group; SI = Symptomatic Group. Note: [a] OR = 2.33 (95% CI = 1.16–4.69).

3.3. Baseline Characteristics Associated with the Severity of COVID-19 among Individuals in the Cohort of Professionals from Belém

The epidemiological characteristics and associated comorbidities were also analyzed in relation to the severity of the disease (between the AS, SCP, and SSP groups), according to the tomographic results of the subjects, and are displayed in Table 3.

Comparison between the AS and SCP groups showed that people with pre-existing comorbidities are 8 times more likely to have lung injuries ($p < 0.001$; OR = 7.9 [95% CI = 3.31–18.98]). Analysis between the SCP and SSP groups demonstrated a 10-fold increase in the risk ($p = 0.016$; OR = 10.26 [95% CI = 1.14–92.25]) of having pulmonary involvement in symptomatic individuals with pre-existing comorbidities. In the comparison between the AS and SCP groups, no asymptomatic individual had two or more comorbidities ($p = 0.012$; OR = 0.48 [95% CI = 0.33 = 0.69]).

Table 3. Epidemiological characteristics of this cohort of professionals exposed to SARS-CoV-2 in the first wave of the pandemic based on COVID-19 severity.

Variable *n* (%)	AS (*n* = 91) *n* (%)	SCP (*n* = 35) *n* (%)	SSP (*n* = 08) *n* (%)	*p*-Value (AS vs. SCP)	*p*-Value (AS vs. SSP)	*p*-Value (SCP vs. SSP)
Age group						
19–34 years	40 (44%)	12 (34.3%)	4 (50%)	$p > 0.05$	$p > 0.05$	$p > 0.05$
35–50 years	43 (47.3%)	18 (51.4%)	4 (50%)			
>50 years	8 (8.8%)	5 (14.3%)	0			
Sex						
Female	67 (73.6%)	21 (60%)	6 (75%)	$p > 0.05$	$p > 0.05$	$p > 0.05$
Male	24 (26.4%)	14 (40%)	2 (25%)			
Presence of pre-existing comorbidities						
No comorbidities	75 (82.4%)	13 (37.1%)	7 (87.5%)	**$p < 0.001$** [a]	$p > 0.05$	**$p = 0.016$** [b]
With comorbidities	16 (17.6%)	22 (62.9%)	1 (12.5%)			
Quantity of comorbidities						
1 comorbidity	16 (100%)	15 (68.2%)	0	**$p = 0.012$** [c]	$p > 0.05$ *	$p > 0.05$
≥2 comorbidities	0	7 (31.8%)	1 (100%)			
Types of comorbidities						
Asthma	4 (25%)	6 (18.75%)	1 (50%)	**$p = 0.018$** [d]	$p > 0.05$ *	$p > 0.05$
Cardiopathies	0	2 (6.25%)	0	**$p = 0.022$** [e]	-	$p > 0.05$
Diabetes mellitus	0	4 (12.5%)	0	**$p = 0.001$** [f]	-	$p > 0.05$ *
Systemic arterial hypertension	5 (31.25%)	6 (18.75%)	0	**$p = 0.038$** [g]	$p > 0.05$ *	$p > 0.05$ *
Overweight and obesity	5 (31.25%)	12 (37.5%)	1 (50%)	**$p < 0.001$** [h]	$p > 0.05$ *	$p > 0.05$
Autoimmune disease	2 (12.5%)	0	0	$p > 0.05$ *	$p > 0.05$ *	-
Pulmonary fibrosis	0	1 (3.125%)	0	$p > 0.05$ *	-	$p > 0.05$ *
Glaucoma	0	1 (3.125%)	0	$p > 0.05$ *	-	$p > 0.05$ *
Profession category						
Administrative	23 (25.3%)	6 (17.1%)	2 (25%)	$p > 0.05$	$p > 0.05$	$p > 0.05$
Healthcare professional	56 (61.5%)	25 (71.4%)	6 (75%)			
General Services	12 (13.2%)	4 (11.4%)	0			

Caption: AS = Asymptomatic Group; SCP = Symptomatic Group with Pulmonary Compromise; SSP = Symptomatic Without Pulmonary Compromise. Note: [a] OR = 7.9 (95% CI = 3.31–18.98); [b] OR = 10.26 (95% CI = 1.14–92.25); [c] OR = 0.48 (95% CI = 0.33 = 0.69); [d] OR = 4.5 (95% CI = 1.18–17.06); [e] OR = 0.26 (95% CI = 0.19–0.35); [f] OR = 0.25 (95% CI = 0.18–0.34); [g] OR = 3.55 (95% CI = 1.01–12.53); [h] OR = 8.97 (95% CI = 2.86–28.06). * Fisher's Exact Test.

The distribution and proportion of the number of individuals with or without comorbidities and lung injury in the AS and SCP groups is shown in Figure 2, which reports that symptoms and lung damage are present only in those with 2 or more comorbidities (100%). In these groups (AS and SCP), the analysis found that asthma ($p = 0.018$; OR = 4.5 [95% CI = 1.18–17.06]), cardiopathies ($p = 0.022$; OR = 0.26 [95% CI = 0.19–0.35]), Diabetes Mellitus ($p = 0.001$; OR = 0.25 [95% CI = 0.18–0.34]), Systemic Arterial Hypertension—SAH ($p = 0.038$; OR = 3.55 [95% CI = 1.01–12.53]), and overweight and obesity ($p < 0.001$; OR = 8.97 [CI 95% = 2.86–28.06]) were significantly correlated; however, for the other investigations between the AS and SSP and SCP and SSP groups and types of comorbidities, no significant associations were observed.

The distribution of individuals with family members who were sick or not and their relationship with kinship and cohabitation were characterized in Table 4 according to the severity groups of the disease. From a family point of view, there was a significant association between two times higher chances of severe illness in family members of study subjects when comparing the SCP with the AS group ($p = 0.022$; OR = 2.86 [95% CI = 1.13–7.24]). Relatives living with the research participant had about seven times more

chances of being part of the SSP group in comparison with being a relative of the AS group ($p = 0.04$; OR = 7.25 [95% CI = 0.81–64.45]).

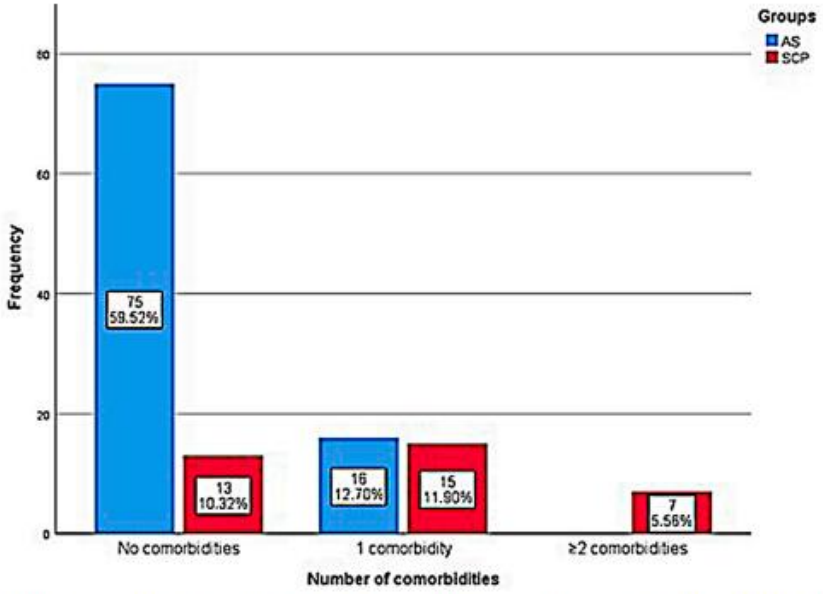


Figure 2. Graphic representation in bars of the absolute and relative number of pre-existing comorbidities in relation to the gravity between the groups of AS and SCP individuals in the cohort.

Table 4. Kinship and housing relationships of family members of individuals exposed to SARS-CoV-2 according to the COVID-19 severity of this cohort of professionals.

Variable *n* (%)	Kinship	AS (*n* = 91) *n* (%)	SCP (*n* = 35) *n* (%)	SSP (*n* = 08) *n* (%)	*p*-Value (AS vs. SCP)	*p*-Value (AS vs. SSP)	*p*-Value (SCP vs. SSP)
Getting Sick							
Relatives who did not get sick		38 (41.76%)	7 (20%)	1 (12.5%)	**$p = 0.022$** [a]	$p > 0.05$	$p > 0.05$
Relatives who became ill		53 (58.24%)	28 (80%)	7 (87.5%)			
Living with the research participant	Blood relatives	24 (45.28%)	18 (64.29%)	6 (85.71%)	$p > 0.05$	**$p = 0.04$** [b]	$p > 0.05$
Not living with the research participant		29 (54.72%)	10 (35.71%)	1 (14.29%)			
Getting sick in the same household							
Relatives who did not get sick	Non-blood relatives	64 (70.3%)	20 (57.1%)	5 (62.5%)	$p > 0.05$	$p > 0.05$	$p > 0.05$
Relatives who did not get sick		27 (29.7%)	15 (42.9%)	3 (37.5%)			

Caption: AS = Asymptomatic Group; SCP = Symptomatic Group with Pulmonary Compromise; SSP = Symptomatic Without Pulmonary Compromise. Note: [a] OR = 2.86 (95% CI = 1.13–7.24); [b] OR = 7.25 (95% CI = 0.81–64.45).

3.4. Genotyping Data for TLR2 SNP rs3804100 Related to Symptomatology and Severity of COVID-19

Genotyping of the *TLR2* SNP rs3804100 is described below, with proportions and associations always being performed by observing the wild-type SNP allele (T) in relation to the mutant (C). For the SNP rs3804100, samples of 214 individuals from the cohort were sequenced, which were analyzed based on the distribution of alleles and genotypes (Figures 3 and 4). No participant in all groups in this research had the C/C genotype (homozygous recessive genotype). Table 5 shows the absolute and relative frequency and

associations of the genotypes and alleles found for the SNP rs3804100 in this cohort in relation to the AS and SI groups.

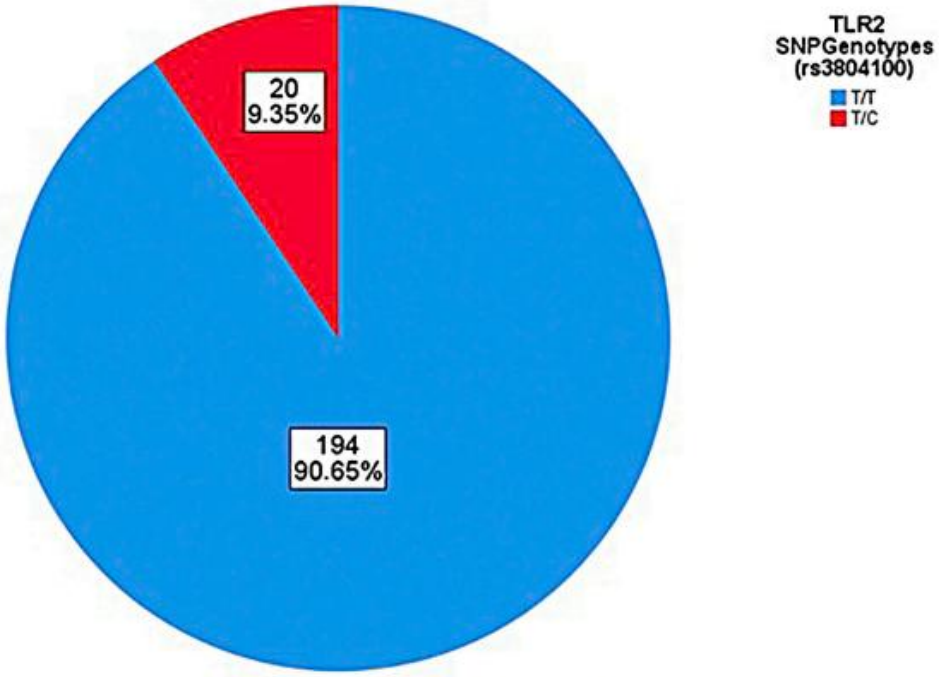


Figure 3. Genotypic distribution of the rs3804100 polymorphism in the sampling.

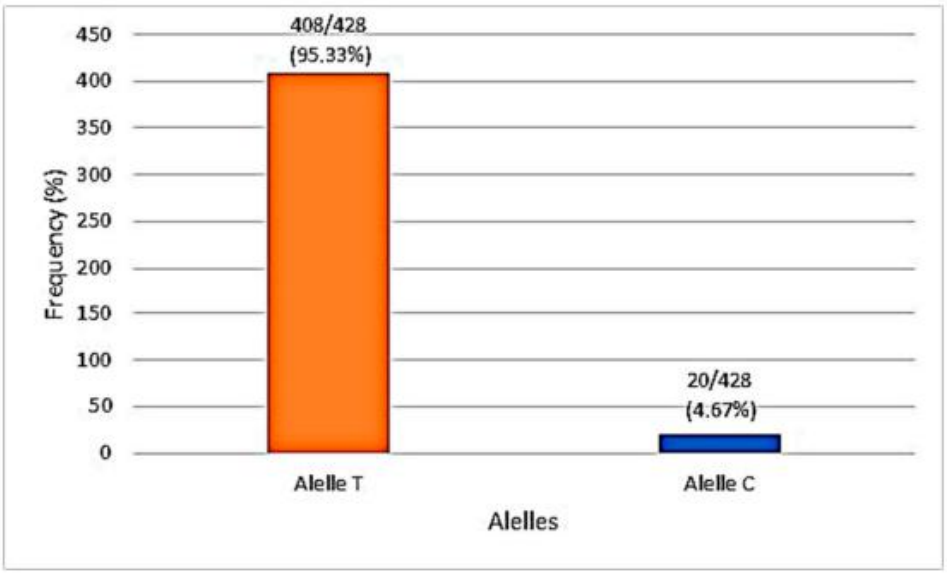


Figure 4. Allelic distribution of the rs3804100 polymorphism in the sampling. Source: Prepared by authors in Microsoft Office Excel 365 (2023).

Table 6 analyzes the absolute and relative frequency of the genotypes and alleles found for the SNP rs3804100 in this cohort in relation to the severity groups of the disease. No participant in the SSP group of this research had the T/C or C/C genotypes.

Table 5. Genotyping of this cohort of professionals for the *TLR2* SNP rs3804100 for symptomatology association.

Genotyping *n* (%)	AS (*n* = 91) *n* (%)	SI (*n* = 123) *n* (%)	*p*-Value
T/T	83 (91.2%)	111 (90.2%)	$p > 0.05$
T/C	8 (8.8%)	12 (9.8%)	
T **(wild allele)**	174 (95.60%)	234 (95.12%)	$p > 0.05$
C	8 (4.4%)	12 (4.88%)	

Caption: AS = Asymptomatic Group; SI = Symptomatic Group.

Table 6. Comparative genotyping of this cohort of professionals for the analyzed *TLR2* SNP related to COVID-19 severity.

Genotyping *n* (%)	AS (*n* = 91) *n* (%)	SCP (*n* = 35) *n* (%)	SSP (*n* = 08) *n* (%)	*p*-Value (AS vs. SCP)	*p*-Value (AS vs. SSP)	*p*-Value (SCP vs. SSP)
T/T	83 (91.2%)	31 (88.6%)	8 (100%)	$p > 0.05$	$p > 0.05$	$p > 0.05$
T/C	8 (8.8%)	4 (11.4%)	0			
T **(wild allele)**	174 (95.6%)	66 (94.3%)	16 (100%)	$p > 0.05$	$p > 0.05$	$p > 0.05$
C	8 (4.4%)	4 (5.7%)	0			

Caption: AS = Asymptomatic Group; SCP = Symptomatic Group with Pulmonary Compromise; SSP = Symptomatic Without Pulmonary Compromise.

3.5. Analysis of the Characteristics of the Non-Synonymous Mutation rs3804100

Functional instability in proteins can occasionally result from structural alterations in amino acid residues due to variations in protein translation. While I-Mutant 2.0 employs a support vector machine (SVM) for the automated prediction of protein stability changes after mutations, PolyPhen-2 works with the sequence, structural, and phylogenetic information of mutations [47].

Using PolyPhen-2, the mutation score, which ranges from 0 to 1, is considered to evaluate if it is harmful. A mutation is considered harmful if the score is close to 1 [48]. I-Mutant 2.0 predicts changes in protein stability (which determines a protein's conformational form) using free energy change values ($\Delta\Delta G$), where a negative value of $\Delta\Delta G$ denotes a protein's declining stability and $\Delta\Delta G > 0$ suggests an increase in stability in the protein's structural conformation. Misfolding, protein breakdown, or abnormal protein aggregation can all result from changes in protein stability [49].

PolyPhen-2 results demonstrated a Benign effect (most likely lacking any phenotypic effect on protein structure or function) prediction for the presence of a mutation in the *TLR2* gene, with a score of 0.002, while I-Mutant 2.0 showed a decreased stability to the protein, with a $\Delta\Delta G$ of -1.35. The 3D model generated by PolyPhen-2 for visualization of structural alterations by the *TLR2* SNP rs3804100 is shown in Figure 5.

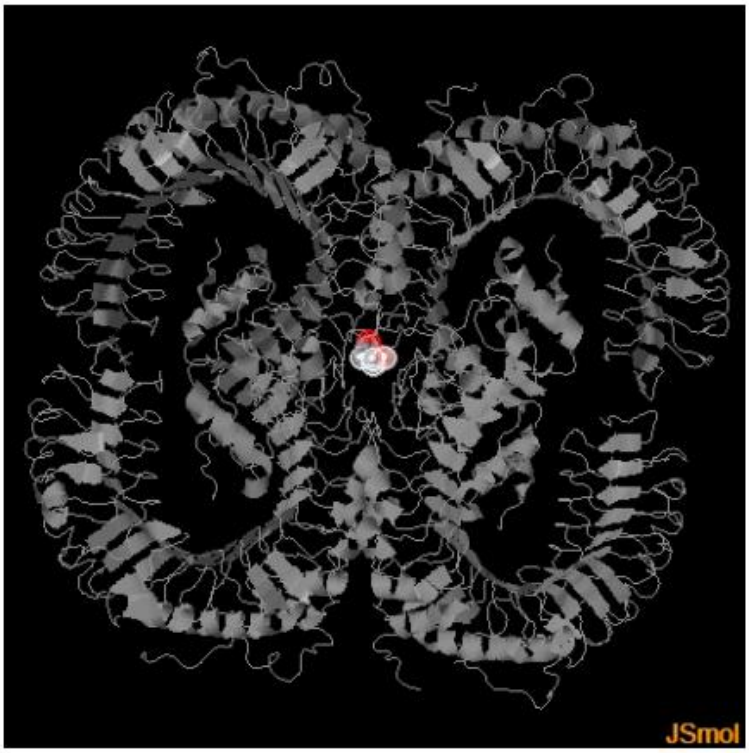


Figure 5. Three-dimensional model of the human TLR2 protein with colored marking in the amino acid change region (Ser450Arg) in its structure due to the presence of the SNP rs3804100. Source: Generated by the PolyPhen-2 server through the Protein Data Bank (PDB)/DSSP Database with UniProt Database Entry O60603 (Human TLR2) and Entry ID 6NIG.

4. Discussion

Risk factors are conditions that increase the likelihood of becoming ill due to a health problem or infectious disease. The most important host risk factors for COVID-19 infection are age, male gender, and comorbid chronic conditions such as hypertension, type 2 diabetes mellitus, obesity, etc. [50]. CCT can be used in symptomatic patients, even with reports of greater sensitivity; however, with limited specificity in relation to virus detection by the standard molecular test [10]. Regarding the use of tomographic results to assess the severity of the disease, this division was chosen to verify whether the group considered more severe had more significant changes, and the SCP group, due to presenting pulmonary involvement greater than or equal to 10%, was considered the most serious parameter [51].

Although it was not possible to characterize bilateral pneumonia with ground-glass manifestation in CCT in these individuals, this analysis of the severity of the disease was carried out in conjunction with the investigation of symptoms suggestive of COVID-19 (particularly to the first pandemic wave) in study professionals directly exposed to the virus. This perspective was considered for the methodology, taking into account the Brazilian reality at the time of scarcity of diagnostic tests and the low availability of individual and collective protective equipment (EPIs/EPCs) in health services (a situation that was still more pronounced in regions far from large urban centers, such as the Amazon Region), and it reduces the risk of bias in relation to the diagnostic question of the research subjects [52–54].

In the present case-control study, the age group did not have a significant relationship between any analyses of the groups involved; however, it must be taken into account that the sample had a work requirement, with most individuals in the 19–50 age category, which represents a percentage of 91.12% of the total number of participants in the cohort, while the age group over 50 years was only 8.88% of the total, with 8.8% in the asymptomatic group and 8.9% in the symptomatic group. The literature reports that SARS-CoV-2 can infect people of any age; however, it is significantly less prevalent and frequently asymptomatic in children and young adults under the age of 14, and, in the first pandemic wave, before the availability of vaccination, the median age of death of the disease was 75 years [55,56].

With regard to gender, no significant statistical associations were observed either, so that all groups had a majority of female members; however, the gender sampling also has an employment relationship, since, in occupational positions in the area of health, the presence of female employees is greater [57]. Nevertheless, the literature indicates a greater risk of severity for men. In an intensive care unit (ICU) cohort from Italy, 82% were men [58]. This disparity in the severity and mortality of COVID-19 may be partially explained by sex-based variations in the expression of the ACE2 receptor and TMPRSS2 [59]. The patient's profession category did not show significant differences between the groups analyzed, regardless of being in the administrative sector, general services, or professional in the health area.

Comorbidities are observed in 24% to 51% of hospitalized patients but in 68% to 72% of ICU patients, indicating that severe COVID-19 is related to them [60]. Among the comorbidities found in this current cohort, Overweight and obesity, and Diabetes Mellitus were significantly associated with symptomatic individuals. Furthermore, it was observed that in the SCP group, 62.9% (22/35) had at least one comorbidity. The most frequent comorbidities were descending: overweight and obesity (38.16%); systemic arterial hypertension and Asthma (19.74% each); Diabetes mellitus (9.21%); heart diseases (5.26%); glaucoma and autoimmune diseases (2.76% each); kidney disease and pulmonary fibrosis (1.32% each). These data are corroborated by the risk factors already observed for COVID-19 in the literature [60,61].

Numerous endocrine changes, including cortisol insufficiency, hypogonadism, and hypothyroidism, play a role in mediating the negative correlation between obesity and the results of COVID-19 [62]. Due to a possible intensification of the inflammatory response to COVID-19 infection and the resulting changes in T cell-mediated immunity, persistent inflammation in obese individuals is hypothesized to be a contributing factor to the observed higher mortality [63–65].

Different regions of the world showed varying prevalence rates of COVID-19 among asthmatics. Although early research from China found a low frequency of asthma among COVID-19 patients [66–69], some US data [70] suggested an increased prevalence of COVID-19 among children with asthma compared to the adult population (14.4% versus 7.8%, respectively), and, in the USA, it is a risk factor for hospitalization [71]. In Europe, there were regional differences in the prevalence of asthma [72–76]. An American study also found that asthma was independently related to a longer intubation time for COVID-19. Obese people with the coexistence of asthma are at much higher risk of a deteriorating illness course from the illness [77].

These variations in epidemiology might be caused by several factors. Various studies use various methods to diagnose asthma, which can lead to overdiagnosis or underdiagnosis. Furthermore, the prevalence of asthma varies by county, which may be the result of variations in environmental exposure, socioeconomic status, and genetic predisposition [78].

The severity and mortality of COVID-19 appear to be correlated with the presence of diabetes mellitus and individual hyperglycemia levels. Poor glycemic control increases the likelihood of SARS-CoV-2 infection in diabetics and increases the death rate, treatment requirements, and hospitalizations [79]. Cardiovascular illnesses such as myocarditis, arrhythmia, cardiogenic shock, heart failure, and other thromboembolic events are caused by the effect of the infection on the cardiovascular system. In addition to hypertensive individ-

uals, the vascular endothelium can activate monocytes, resulting in almost uncontrollable cytokine production that may have a connection to COVID-19 [80].

In this present study, it was also possible to verify that, although not statistically significant, in the symptomatic group, 67.48% of the consanguineous relatives of these individuals became ill, and, among these relatives, 69.88% of the household members became ill, with about seven times greater chance of becoming ill compared to the extradomiciliary. These data suggest that the illness of family members was influenced by contact in the same house with health professionals. These data suggest that family illness is linked to an association of genetic and environmental factors. Regarding ambiental aspects, it is known that to prevent the spread of SARS-CoV-2 infected viral particles and aerosols, it is recommended to avoid interpersonal contact, wash your hands after touching any physical surfaces in anthropogenic settings, and adhere strictly to respiratory etiquette rules [81]. When dealing with people who have a suspected or proven infection, quarantine and confinement are also recommended to prevent disease transmission [82,83].

Subjects without associated risk factors also experience symptoms and worsening of the disease, indicating that genetic factors can also affect the condition and that SNPs in specific genes can affect the variation in the incidence and severity of the disease by dysregulating the host immune response. The innate immune response originates from viral entry into the human cell. When entering the body, the first form of contact between SARS-CoV-2 and humans is through the PRRs of antigen-presenting cells (APCs), such as monocytes and dendritic cells (DCs), which recognize the PAMPs and DAMPs of the virus. In this context, the main PRRs responsible for detecting this virus are TLRs, retinoic acid-inducible I-like receptors (RLRs), nucleotide-binding oligomerization domain-like receptors (NLRs), and absent in melanoma 2 (AIM-2) [84,85]. These proteins induce signaling pathways, such as interferon receptor genes alpha (*IFNAR*) *1* and *IFNAR2*, which produce interferon (IFN)-I and IFN-III, mediators already considered potential inhibitors of the virus in vitro [86].

TLR2 is an extracellular sensor present in the membrane of immune cells, such as macrophages and DCs, that, through pairs or dimers, performs its function of detecting the pathogen through the innate immune response. The ability of TLR2 to differentiate between diacyl and triacyl lipopeptides is achieved by homodimerization or heterodimerization with TLR1 and TLR6, respectively [87]. It can recognize a number of bacterial, viral, fungal, and parasitic agents [88].

In the case of SARS-CoV-2 infection, in vitro analysis in mice demonstrated that anti-TLR2 therapy has positive impacts on the survival of these animals. Disruption of TLR2 did not affect SARS-CoV-2-mediated type I IFN-α (a type of IFN-I) production but did significantly alter IL-6 production (a marker of a poor prognosis) [18]. A study by Van Der Sluis et al. (2022) confirmed that plasmacytoid DCs (pDCs) produce IL-6 in response to TLR2 and TLR2/6-mediated sensing of SARS-CoV-2 glycoproteins [89]. Furthermore, there is evidence that activation of the NF-kB inflammatory pathway in COVID-19 (a disease aggravating factor) by protein S is dependent on TLR2 signaling [90]. Therefore, the induction of TLR2 by the human innate immune system probably has more unfavorable outcomes than a protective effect in COVID-19. However, this study focuses on specific *TLR2* genes rather than *IFN* or *IL-6* genes because we place an emphasis on studying the analysis of genes that encode immune cell receptors.

The human *TLR2* gene is located on chromosome 4q31.3 [91]. This gene consists of 3 exons (coding regions) [92]. There are several SNPs in the *TLR2* gene, including the SNP rs3804100. The SNP rs3804100, also known as 1350T/C, is a non-synonymous mutation of the missense type; that is, it leads to a change in the codon from an amino acid in Serine (Ser) to Arginine (Arg) at residue 450. This SNP was first described in a study by Butty et al. (2008) in an analysis of their susceptibility to type 1 diabetes, in which this polymorphism did not have significant differences in frequencies between diabetic progressors and non-progressors [93].

The frequency of the mutant allele (C) of this SNP found in the present study for the Amazonian population, more precisely Belém city, was 9.35%, a slightly different percentage from that found reported in the NCBI for the Latino population. americana, estimated at 7.52% [37]. An immunogenetic study with the same SNP in the region of Brazil, the State of Pará, but under analysis for patients with leprosy, found a frequency of the mutant allele of approximately 4.1% and also did not find its homozygous recessive genotype (C/C) [94].

This SNP was associated with a decrease in exonic splicing site domains, however, with no defined function in *TLR2* gene expression yet [95–97]. Therefore, this study proposes an evaluation of the biological and conformational protein effects of this SNP in *TLR2* through computational prediction. The results demonstrated that this SNP is benign in individuals and causes a decrease in the stability of the TLR2 protein, which may indicate an inverse relationship to the original function of TLR2 in a given population with a disease, generated by a possible functional instability in TLR2 that is not yet fully understood.

The presence of *TLR2* SNPs in each population can functionally affect the production of the receptor and therefore impact the generation of an effective immune response [21]. Despite having been predicted in this present work about the benign nature and the reduction in the protein stability of TLR2 related to the presence of the SNP rs3804100 and the literary scenario more related to the patient's worse prognosis when this gene is more expressed, in the case of the Amazonian population, significant associations for any investigated COVID-19 phenotype were not observed. This may simply indicate that this gene region has different genetic variability in people from Belém in relation to other populations, due to gene-gene and gene-environment interactions [27].

Understanding the genetic basis of the host refers to a possible correlation with the manifestation of diseases, as genetic expression in DNA is capable of deregulating in certain people the construction of an effective response against possible health problems [98]. So, this present study is a pioneer in the search for associations between this *TLR2* SNP and the severity and manifestations of COVID-19 in the Brazilian population and the second to investigate these associations with the disease worldwide. In addition to that, it sought to predict biological functions and protein relationships related to the presence of this SNP through bioinformatics analysis.

The study by Salamaikina et al. (2022) with a Russian population was the only study to date investigating this SNP and COVID-19; however, it had a different objective from this present study, which was to compare the allele frequencies of genetic polymorphisms in *TLR* genes in samples obtained from individuals with pneumonia in the period before the appearance of SARS-CoV-2 and during the COVID-19 pandemic in this same population and with international public data from the European population. In contrast to the results of this present research, they concluded that Russian individuals with the mutant allele (C) of the SNP rs3804100 have a 0.32-fold greater vulnerability to COVID-19 pneumonia compared to the group of healthy people [99].

The results of the association of SNPs can vary from one population to another due to the background phenomenon that, together with the environmental circumstances of each distinct population group, can determine the immunogenetic heterogeneity [100,101]. In this sense, further research is needed to better visualize epidemiological and genetic parameters and their consequences in COVID-19. Moreover, more studies are required to complement the information obtained, and more SNPs of the *TLR2* gene must be looked at in relation to the clinical manifestations and severity in this group or the same SNP in cohorts with a larger sample number.

Being a retrospective study, one of the drawbacks of this study is that a COVID-19 test could not be conducted in all of the cohorts, as we were unable to reach the patients when they were actively working on the first wave in the health centers. Additionally, during the first wave, many healthcare providers did not perform tests for the disease, as they were hard to come by and in short supply [102]. Since it was not possible to access tomography data for the entire cohort, the dyspnea symptom was excluded from a possible

severity classification because it was impossible to define it as a COVID-19 symptom or a psychological factor. Instead, it was only considered at the end of the analyses based on the participants' reports from the questionnaires. Additionally, it has memory constraints since the data we evaluated was based on a questionnaire. Our research, however, is restricted to the time of the first wave, which we believe individuals may continue to remember since it was such a horrific and significant time in recent history [103,104].

5. Conclusions

This work looked for associations between risk factors and the *TLR2* SNP rs3804100 in Health Institution workers and COVID-19. In this way, a relationship was observed between the occurrence of symptoms (SI) and comorbidity, with a higher prevalence of overweight and obese individuals. Similarly, a relationship was established between the occurrence of symptoms and pulmonary involvement and having two or more comorbidities, with a higher prevalence of comorbidities such as diabetes mellitus, overweight and obesity, systemic arterial hypertension, heart disease, and asthma.

The presence of comorbidity was also more present in the group of symptomatic subjects with pulmonary involvement than in those with asymptomatic ones, and it also predominated in those with lung injury to the detriment of those without lung injury. There was an observation that 100% of the symptomatic professionals with pulmonary involvement in the cohort had $\geq$2 comorbidities.

There was an association between sick blood relatives living in the same household as those with symptoms and a relationship between illness in blood relatives and symptomatology with pulmonary involvement (irrespective of the household). Furthermore, consanguineous relatives of cohort subjects without pulmonary involvement became significantly ill at the same home.

In the genotype analysis of the *TLR2* SNP rs3804100 (−1350 T/C), the homozygous recessive (C/C) genotype was not found, and the frequency of the mutant allele (C) was less than 10% in the cohort. No significant associations were found for this SNP in this cohort of Belém Health Institutions professionals in the analysis for symptomatology or severity. Furthermore, bioinformatics analysis suggested that the presence of this SNP is benign in individuals and causes a decrease in the stability of the TLR2 protein.

Therefore, the observed data on epidemiological and immunogenetic markers can help the scientific community and medicine seek new forms of containment (especially from the point of view of professionals in the clinical-hospital structure) and combat the disease, such as those focused on anti-TLR2 therapy.

Author Contributions: Conceptualization, M.J.A.S., K.V.B.L. and L.N.G.C.L.; methodology, M.J.A.S., C.S.S. and L.N.G.C.L.; validation, P.A.S.d.S., S.M.M.C., K.V.B.L. and L.N.G.C.L.; formal analysis, M.J.A.S. and C.S.S.; investigation, P.A.S.d.S., R.L.M., C.S.S., E.P.d.C.G. and J.G.C.; resources, E.P.d.C.G., J.G.C., K.V.B.L. and L.N.G.C.L., writing—original draft preparation, M.J.A.S., K.V.B.L., C.S.S., R.L.M. and S.M.M.C.; writing—review and editing, K.V.B.L. and L.N.G.C.L.; visualization, S.M.M.C., K.V.B.L. and L.N.G.C.L.; supervision, K.V.B.L. and L.N.G.C.L.; project administration, L.N.G.C.L.; funding acquisition, K.V.B.L. and L.N.G.C.L. All authors have read and agreed to the published version of the manuscript.

Funding: Marcos Jessé Abrahão Silva was funded by Bolsa CAPES—Finance Code 88887.888066/2023-00, and the APC was funded by the Evandro Chagas Institute (IEC).

Institutional Review Board Statement: Approved by the Research Ethics Committee—CEP of the State University of Pará (UEPA), Brazil, under opinion number 6.124.862.

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: The original contributions to this study are included in the article. Further inquiries can be directed to the corresponding authors.

Acknowledgments: The authors are thankful to the study subjects and the Graduate Program in Epidemiology and Health Surveillance (PPGEVS), Evandro Chagas Institute (IEC), Brazil.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, X.; Wang, W.; Zhao, X.; Zai, J.; Zhao, Q.; Li, Y.; Chaillon, A. Transmission Dynamics and Evolutionary History of 2019-nCoV. *J. Med. Virol.* **2020**, *92*, 501–511. [CrossRef]
2. Harrison, A.G.; Lin, T.; Wang, P. Mechanisms of SARS-CoV-2 Transmission and Pathogenesis. *Trends Immunol.* **2020**, *41*, 1100–1115. [CrossRef]
3. World Health Organization. *Coronavirus Disease 2019 (COVID-19): Situation Report, 72*; World Health Organization: Geneva, Switzerland, 2020.
4. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. Available online: https://covid19.who.int (accessed on 14 March 2023).
5. Zeiser, F.A.; Donida, B.; da Costa, C.A.; de Oliveira Ramos, G.; Scherer, J.N.; Barcellos, N.T.; Alegretti, A.P.; Ikeda, M.L.R.; Müller, A.P.W.C.; Bohn, H.C.; et al. First and Second COVID-19 Waves in Brazil: A Cross-Sectional Study of Patients' Characteristics Related to Hospitalization and in-Hospital Mortality. *Lancet Reg. Health Am.* **2022**, *6*, 100107. [CrossRef]
6. Zimmermann, I.R.; Sanchez, M.N.; Frio, G.S.; Alves, L.C.; Pereira, C.C.d.A.; Lima, R.T.d.S.; Machado, C.; Santos, L.M.P.; Silva, E.N. da Trends in COVID-19 Case-Fatality Rates in Brazilian Public Hospitals: A Longitudinal Cohort of 398,063 Hospital Admissions from 1st March to 3rd October 2020. *PLoS ONE* **2021**, *16*, e0254633. [CrossRef]
7. de Albuquerque Costa, W.; de Campos Carvalho, N.; Santana, V.R.; Coelho, P.A.B.; Moreira, A.M.; do Nascimento, M.S. Políticas Públicas e as lições preliminares da COVID-19 na Atenção Primária à Saúde da Ceilândia-DF. *Comun. Ciências Saúde* **2022**, *33*, 129–142. [CrossRef]
8. To, K.K.-W.; Sridhar, S.; Chiu, K.H.-Y.; Hung, D.L.-L.; Li, X.; Hung, I.F.-N.; Tam, A.R.; Chung, T.W.-H.; Chan, J.F.-W.; Zhang, A.J.-X.; et al. Lessons Learned 1 Year after SARS-CoV-2 Emergence Leading to COVID-19 Pandemic. *Emerg. Microbes Infect.* **2021**, *10*, 507–535. [CrossRef] [PubMed]
9. Araujo-Filho, J.d.A.B.; Sawamura, M.V.Y.; Costa, A.N.; Cerri, G.G.; Nomura, C.H. Pneumonia por COVID-19: Qual o papel da imagem no diagnóstico? *J. Bras. Pneumol.* **2020**, *46*, e20200114. [CrossRef] [PubMed]
10. Kanne, J.P.; Little, B.P.; Chung, J.H.; Elicker, B.M.; Ketai, L.H. Essentials for Radiologists on COVID-19: An Update-Radiology Scientific Expert Panel. *Radiology* **2020**, *296*, E113–E114. [CrossRef] [PubMed]
11. Simpson, S.; Kay, F.U.; Abbara, S.; Bhalla, S.; Chung, J.H.; Chung, M.; Henry, T.S.; Kanne, J.P.; Kligerman, S.; Ko, J.P.; et al. Radiological Society of North America Expert Consensus Statement on Reporting Chest CT Findings Related to COVID-19. Endorsed by the Society of Thoracic Radiology, the American College of Radiology, and RSNA—Secondary Publication. *J. Thorac. Imaging* **2020**, *35*, 219–227. [CrossRef]
12. BRASIL. *Boletim Epidemiológico No 109—Boletim COE Coronavírus—Português (Brasil)*; Boletim Epidemiológico 109–COE COVID-26 de abril de 2022; Ministério da Saúde, Centro de Operações de Emergência em Saúde Pública/Doença Pelo Coronavírus: Brasília, Brazil, 2022.
13. World Health Organization. COVID 19 Public Health Emergency of International Concern (PHEIC) Global Research and Innovation Forum: Towards a Research Roadmap. *Glob. Res. Collab. Infect. Dis. Prep.* **2020**, *11*, e02007.
14. Hou, Y.; Zhao, J.; Martin, W.; Kallianpur, A.; Chung, M.K.; Jehi, L.; Sharifi, N.; Erzurum, S.; Eng, C.; Cheng, F. New Insights into Genetic Susceptibility of COVID-19: An ACE2 and TMPRSS2 Polymorphism Analysis. *BMC Med.* **2020**, *18*, 216. [CrossRef] [PubMed]
15. Ou, X.; Liu, Y.; Lei, X.; Li, P.; Mi, D.; Ren, L.; Guo, L.; Guo, R.; Chen, T.; Hu, J.; et al. Characterization of Spike Glycoprotein of SARS-CoV-2 on Virus Entry and Its Immune Cross-Reactivity with SARS-CoV. *Nat. Commun.* **2020**, *11*, 1620. [CrossRef] [PubMed]
16. Rokni, M.; Ghasemi, V.; Tavakoli, Z. Immune Responses and Pathogenesis of SARS-CoV-2 during an Outbreak in Iran: Comparison with SARS and MERS. *Rev. Med. Virol.* **2020**, *30*, e2107. [CrossRef] [PubMed]
17. Nemazee, D.; Gavin, A.; Hoebe, K.; Beutler, B. Toll-like Receptors and Antibody Responses. *Nature* **2006**, *441*, E4. [CrossRef] [PubMed]
18. Zheng, M.; Karki, R.; Williams, E.P.; Yang, D.; Fitzpatrick, E.; Vogel, P.; Jonsson, C.B.; Kanneganti, T.-D. TLR2 Senses the SARS-CoV-2 Envelope Protein to Produce Inflammatory Cytokines. *Nat. Immunol.* **2021**, *22*, 829–838. [CrossRef]
19. Root-Bernstein, R. Innate Receptor Activation Patterns Involving TLR and NLR Synergisms in COVID-19, ALI/ARDS and Sepsis Cytokine Storms: A Review and Model Making Novel Predictions and Therapeutic Suggestions. *Int. J. Mol. Sci.* **2021**, *22*, 2108. [CrossRef]
20. Caetano, A.R. SNP Markers: Basic Concepts, Applications in Animal Breeding and Management and Perspectives for the Future. *R. Bras. Zootec.* **2009**, *38*, 64–71. [CrossRef]
21. Silva, M.J.A.; Lima, M.B.M.; Lima, K.V.B.; Lima, L.N.G.C. The Relationship of TLR2 Polymorphisms with Infectious Diseases. *Annu. Res. Rev. Biol.* **2021**, *13*, 57–73. [CrossRef]
22. Silva, M.J.A.; Santana, D.S.; de Oliveira, L.G.; Monteiro, E.O.L.; Lima, L.N.G.C. The Relationship between 896A/G (Rs4986790) Polymorphism of TLR4 and Infectious Diseases: A Meta-Analysis. *Front. Genet.* **2022**, *13*, 1045725. [CrossRef]
23. Silva, M.J.A.; Silva, C.S.; da Silva Vieira, M.C.; dos Santos, P.A.S.; Frota, C.C.; Lima, K.V.B.; Lima, L.N.G.C. The Relationship between TLR3 Rs3775291 Polymorphism and Infectious Diseases: A Meta-Analysis of Case-Control Studies. *Genes* **2023**, *14*, 1311. [CrossRef]

24. Albiger, B.; Dahlberg, S.; Henriques-Normark, B.; Normark, S. Role of the Innate Immune System in Host Defence against Bacterial Infections: Focus on the Toll-like Receptors. *J. Intern. Med.* **2007**, *261*, 511–528. [CrossRef] [PubMed]
25. Wang, T.; Lafuse, W.P.; Zwilling, B.S. Regulation of Toll-like Receptor 2 Expression by Macrophages Following Mycobacterium Avium Infection. *J. Immunol.* **2000**, *165*, 6308–6313. [CrossRef] [PubMed]
26. Yadav, M.; Schorey, J.S. The β-Glucan Receptor Dectin-1 Functions Together with TLR2 to Mediate Macrophage Activation by Mycobacteria. *Blood* **2006**, *108*, 3168–3175. [CrossRef] [PubMed]
27. Casanova, J.-L.; Abel, L. Human Genetics of Infectious Diseases: A Unified Theory. *EMBO J.* **2007**, *26*, 915–922. [CrossRef]
28. Von Elm, E.; Altman, D.G.; Egger, M.; Pocock, S.J.; Gøtzsche, P.C.; Vandenbroucke, J.P.; Initiative, S. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. *Int. J. Surg.* **2014**, *12*, 1495–1499. [CrossRef]
29. Helsinki Declaration. *World Medical Association's Declaration of Helsinki (with Later Amendments): Ethical Principles for Medical Research Involving Human Subjects*; World Medical Association Helsinki: Helsinki, Finland, 1964.
30. Brasil, M.S. Resolução 466/12 Do Conselho Nacional de Saúde. In *Sobre Diretrizes e Normas Regulamentadoras de Pesquisa Envolvendo Seres Humanos*; Conselho Nacional de Saúde: Brasília, Brazil, 2012.
31. Quigley, A.L.; Stone, H.; Nguyen, P.Y.; Chughtai, A.A.; MacIntyre, C.R. Estimating the Burden of COVID-19 on the Australian Healthcare Workers and Health System during the First Six Months of the Pandemic. *Int. J. Nurs. Stud.* **2021**, *114*, 103811. [CrossRef]
32. World Health Organization. *COVID-19: Occupational Health and Safety for Health Workers: Interim Guidance, 2 February 2021*; World Health Organization: Geneva, Switzerland, 2021.
33. Carvajal, D.L.M.; Borges, M.A.S.B.; Zara, A.L.d.S.A.; Pereira, M.S.; da Cunha, L.P.; Turchi, M.D. Perfil Clínico Epidemiológico E Desfechos da Primeira Onda de COVID-19 em Anápolis, Goiás. *Braz. J. Infect. Dis.* **2022**, *26*, 102048. [CrossRef]
34. López-Juárez, P.; Serrano-Oviedo, L.; Pérez-Ortiz, J.M.; García-Jabalera, I.; Bejarano-Ramírez, N.; Gómez-Romero, F.J.; Muñoz-Rodríguez, J.R.; Redondo-Calvo, F.J. Comparative study of the COVID-19 admissions between first and second wave in a cohort of 1,235 patients. *Rev. Esp. Quim.* **2021**, *34*, 387–389. [CrossRef]
35. Choi, H.; Qi, X.; Yoon, S.H.; Park, S.J.; Lee, K.H.; Kim, J.Y.; Lee, Y.K.; Ko, H.; Kim, K.H.; Park, C.M.; et al. Extension of Coronavirus Disease 2019 on Chest CT and Implications for Chest Radiographic Interpretation. *Radiol. Cardiothorac. Imaging* **2020**, *2*, e200107. [CrossRef]
36. UNA-SUS Organização Mundial de Saúde Declara Pandemia Do Novo Coronavírus. Available online: https://www.unasus.gov.br/noticia/organizacao-mundial-de-saude-declara-pandemia-de-coronavirus (accessed on 6 September 2022).
37. Sayers, E.W.; Bolton, E.E.; Brister, J.R.; Canese, K.; Chan, J.; Comeau, D.C.; Farrell, C.M.; Feldgarden, M.; Fine, A.M.; Funk, K. Database Resources of the National Center for Biotechnology Information in 2023. *Nucleic Acids Res.* **2023**, *51*, D29. [CrossRef]
38. Untergasser, A.; Nijveen, H.; Rao, X.; Bisseling, T.; Geurts, R.; Leunissen, J.A.M. Primer3Plus, an Enhanced Web Interface to Primer3. *Nucleic Acids Res.* **2007**, *35*, W71–W74. [CrossRef]
39. Benson, D.A.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Rapp, B.A.; Wheeler, D.L. GenBank. *Nucleic Acids Res.* **2000**, *28*, 15–18. [CrossRef] [PubMed]
40. Lisby, G. Application of Nucleic Acid Amplification in Clinical Microbiology. *Mol. Biotechnol.* **1999**, *12*, 75–100. [CrossRef]
41. Hall, T.A. BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. In *Nucleic Acids Symposium Series*; Information Retrieval Ltd.: London, UK, 1999; Volume 41, pp. 95–98.
42. Chen, B.; Cole, J.W.; Grond-Ginsbach, C. Departure from Hardy Weinberg Equilibrium and Genotyping Error. *Front. Genet.* **2017**, *8*, 167. [CrossRef]
43. Schneider, S.; Roessli, D.; Excoffier, L. Arlequin: A Software for Population Genetics Data Analysis. *User Man. Ver.* **2000**, *2*, 2496–2497.
44. Kang, H. Sample Size Determination and Power Analysis Using the G*Power Software. *J. Educ. Eval. Health Prof.* **2021**, *18*, 17. [CrossRef] [PubMed]
45. Ruch, P.; Teodoro, D.; Consortium, U. UniProt: The Universal Protein Knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480–D489.
46. Berman, H.; Henrick, K.; Nakamura, H. Announcing the Worldwide Protein Data Bank. *Nat. Struct. Mol. Biol.* **2003**, *10*, 980. [CrossRef]
47. Doss, C.G.P.; Rajith, B.; Garwasis, N.; Mathew, P.R.; Raju, A.S.; Apoorva, K.; William, D.; Sadhana, N.R.; Himani, T.; Dike, I. Screening of Mutations Affecting Protein Stability and Dynamics of FGFR1—A Simulation Analysis. *Appl. Transl. Genom.* **2012**, *1*, 37–43. [CrossRef]
48. Adzhubei, I.; Jordan, D.M.; Sunyaev, S.R. Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2. *Curr. Protoc. Hum. Genet.* **2013**, *76*, 7.20.1–7.20.41. [CrossRef]
49. Capriotti, E.; Fariselli, P.; Casadio, R. I-Mutant2. 0: Predicting Stability Changes upon Mutation from the Protein Sequence or Structure. *Nucleic Acids Res.* **2005**, *33*, W306–W310. [CrossRef] [PubMed]
50. Rashedi, J.; Mahdavi Poor, B.; Asgharzadeh, V.; Pourostadi, M.; Samadi Kafil, H.; Vegari, A.; Tayebi-Khosroshahi, H.; Asgharzadeh, M. Risk Factors for COVID-19. *Infez. Med.* **2020**, *28*, 469–474. [PubMed]
51. Berlin, D.A.; Gulick, R.M.; Martinez, F.J. Severe Covid-19. *N. Engl. J. Med.* **2020**, *383*, 2451–2460. [CrossRef] [PubMed]

52. Lv, M.; Wang, M.; Yang, N.; Luo, X.; Li, W.; Chen, X.; Liu, Y.; Ren, M.; Zhang, X.; Wang, L.; et al. Chest Computed Tomography for the Diagnosis of Patients with Coronavirus Disease 2019 (COVID-19): A Rapid Review and Meta-Analysis. *Ann. Transl. Med.* **2020**, *8*, 622. [CrossRef]
53. Ai, T.; Yang, Z.; Hou, H.; Zhan, C.; Chen, C.; Lv, W.; Tao, Q.; Sun, Z.; Xia, L. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology* **2020**, *296*, E32–E40. [CrossRef]
54. Fang, Y.; Zhang, H.; Xie, J.; Lin, M.; Ying, L.; Pang, P.; Ji, W. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* **2020**, *296*, E115–E117. [CrossRef]
55. Lai, C.-C.; Liu, Y.H.; Wang, C.-Y.; Wang, Y.-H.; Hsueh, S.-C.; Yen, M.-Y.; Ko, W.-C.; Hsueh, P.-R. Asymptomatic Carrier State, Acute Respiratory Disease, and Pneumonia Due to Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): Facts and Myths. *J. Microbiol. Immunol. Infect.* **2020**, *53*, 404–412. [CrossRef]
56. Wang, L.; Wang, Y.; Ye, D.; Liu, Q. Review of the 2019 Novel Coronavirus (SARS-CoV-2) Based on Current Evidence. *Int. J. Antimicrob. Agents* **2020**, *55*, 105948. [CrossRef]
57. International Labour Office. *Women at Work: Trends 2016*; ILO: Geneva, Switzerland, 2016; ISBN 92-2-130796-4.
58. Grasselli, G.; Zangrillo, A.; Zanella, A.; Antonelli, M.; Cabrini, L.; Castelli, A.; Cereda, D.; Coluccello, A.; Foti, G.; Fumagalli, R.; et al. Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy. *JAMA* **2020**, *323*, 1574. [CrossRef]
59. Mukherjee, S.; Pahan, K. Is COVID-19 Gender-Sensitive? *J. Neuroimmune Pharmacol.* **2021**, *16*, 38–47. [CrossRef]
60. Yang, J.; Zheng, Y.; Gou, X.; Pu, K.; Chen, Z.; Guo, Q.; Ji, R.; Wang, H.; Wang, Y.; Zhou, Y. Prevalence of Comorbidities and Its Effects in Patients Infected with SARS-CoV-2: A Systematic Review and Meta-Analysis. *Int. J. Infect. Dis.* **2020**, *94*, 91–95. [CrossRef] [PubMed]
61. Russell, C.D.; Lone, N.I.; Baillie, J.K. Comorbidities, Multimorbidity and COVID-19. *Nat. Med.* **2023**, *29*, 334–343. [CrossRef] [PubMed]
62. Stern, J.H.; Rutkowski, J.M.; Scherer, P.E. Adiponectin, Leptin, and Fatty Acids in the Maintenance of Metabolic Homeostasis through Adipose Tissue Crosstalk. *Cell Metab.* **2016**, *23*, 770–784. [CrossRef] [PubMed]
63. Juno, J.A.; Tan, H.-X.; Lee, W.S.; Reynaldi, A.; Kelly, H.G.; Wragg, K.; Esterbauer, R.; Kent, H.E.; Batten, C.J.; Mordant, F.L.; et al. Humoral and Circulating Follicular Helper T Cell Responses in Recovered Patients with COVID-19. *Nat. Med.* **2020**, *26*, 1428–1434. [CrossRef]
64. Misumi, I.; Starmer, J.; Uchimura, T.; Beck, M.A.; Magnuson, T.; Whitmire, J.K. Obesity Expands a Distinct Population of T Cells in Adipose Tissue and Increases Vulnerability to Infection. *Cell Rep.* **2019**, *27*, 514–524.e5. [CrossRef] [PubMed]
65. Yu, L.; Zhang, X.; Ye, S.; Lian, H.; Wang, H.; Ye, J. Obesity and COVID-19: Mechanistic Insights from Adipose Tissue. *J. Clin. Endocrinol. Metab.* **2022**, *107*, 1799–1811. [CrossRef]
66. Li, X.; Xu, S.; Yu, M.; Wang, K.; Tao, Y.; Zhou, Y.; Shi, J.; Zhou, M.; Wu, B.; Yang, Z.; et al. Risk Factors for Severity and Mortality in Adult COVID-19 Inpatients in Wuhan. *J. Allergy Clin. Immunol.* **2020**, *146*, 110–118. [CrossRef]
67. Wu, Z.; McGoogan, J.M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* **2020**, *323*, 1239. [CrossRef]
68. Zhang, J.; Cao, Y.; Dong, X.; Wang, B.; Liao, M.; Lin, J.; Yan, Y.; Akdis, C.A.; Gao, Y. Distinct Characteristics of COVID-19 Patients with Initial rRT-PCR-positive and rRT-PCR-negative Results for SARS-CoV-2. *Allergy* **2020**, *75*, 1809–1812. [CrossRef]
69. Zhang, J.-J.; Dong, X.; Cao, Y.-Y.; Yuan, Y.-D.; Yang, Y.-B.; Yan, Y.-Q.; Akdis, C.A.; Gao, Y.-D. Clinical Characteristics of 140 Patients Infected with SARS-CoV-2 in Wuhan, China. *Allergy* **2020**, *75*, 1730–1741. [CrossRef]
70. Garg, S.; Kim, L.; Whitaker, M.; O'Halloran, A.; Cummings, C.; Holstein, R.; Prill, M.; Chai, S.J.; Kirley, P.D.; Alden, N.B.; et al. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019—COVID-NET, 14 States, March 1–30, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **2020**, *69*, 458–464. [CrossRef] [PubMed]
71. Gaietto, K.; Freeman, M.C.; DiCicco, L.A.; Rauenswinter, S.; Squire, J.R.; Aldewereld, Z.; Iagnemma, J.; Campfield, B.T.; Wolfson, D.; Kazmerski, T.M.; et al. Asthma as a Risk Factor for Hospitalization in Children with COVID-19: A Nested Case-control Study. *Pediatr. Allergy Immunol.* **2022**, *33*, e13696. [CrossRef] [PubMed]
72. Bloom, C.I.; Cullinan, P.; Wedzicha, J.A. Asthma Phenotypes and COVID-19 Risk: A Population-Based Observational Study. *Am. J. Respir. Crit. Care Med.* **2022**, *205*, 36–45. [CrossRef] [PubMed]
73. Caminati, M.; Lombardi, C.; Micheletto, C.; Roca, E.; Bigni, B.; Furci, F.; Girelli, D.; Senna, G.; Crisafulli, E. Asthmatic Patients in COVID-19 Outbreak: Few Cases despite Many Cases. *J. Allergy Clin. Immunol.* **2020**, *146*, 541–542. [CrossRef] [PubMed]
74. Gémes, K.; Talbäck, M.; Modig, K.; Ahlbom, A.; Berglund, A.; Feychting, M.; Matthews, A.A. Burden and Prevalence of Prognostic Factors for Severe COVID-19 in Sweden. *Eur. J. Epidemiol.* **2020**, *35*, 401–409. [CrossRef]
75. Khawaja, A.P.; Warwick, A.N.; Hysi, P.G.; Kastner, A.; Dick, A.; Khaw, P.T.; Tufail, A.; Foster, P.J.; Khaw, K.-T. *Associations with Covid-19 Hospitalisation amongst 406,793 Adults: The UK Biobank Prospective Cohort Study*; Epidemiology: London, UK, 2020.
76. Ren, J.; Pang, W.; Luo, Y.; Cheng, D.; Qiu, K.; Rao, Y.; Zheng, Y.; Dong, Y.; Peng, J.; Hu, Y.; et al. Impact of Allergic Rhinitis and Asthma on COVID-19 Infection, Hospitalization, and Mortality. *J. Allergy Clin. Immunol. Pract.* **2022**, *10*, 124–133. [CrossRef]
77. Mahdavinia, M.; Foster, K.J.; Jauregui, E.; Moore, D.; Adnan, D.; Andy-Nweye, A.B.; Khan, S.; Bishehsari, F. Asthma Prolongs Intubation in COVID-19. *J. Allergy Clin. Immunol. Pract.* **2020**, *8*, 2388–2391. [CrossRef]
78. Adir, Y.; Saliba, W.; Beurnier, A.; Humbert, M. Asthma and COVID-19: An Update. *Eur. Respir. Rev.* **2021**, *30*, 210152. [CrossRef]

79. Sharma, P.; Behl, T.; Sharma, N.; Singh, S.; Grewal, A.S.; Albarrati, A.; Albratty, M.; Meraya, A.M.; Bungau, S. COVID-19 and Diabetes: Association Intensify Risk Factors for Morbidity and Mortality. *Biomed. Pharmacother.* **2022**, *151*, 113089. [CrossRef]
80. Koyama, A.K.; Imperatore, G.; Rolka, D.B.; Lundeen, E.; Rutkowski, R.E.; Jackson, S.L.; He, S.; Kuklina, E.V.; Park, S.; Pavkov, M.E. Risk of Cardiovascular Disease After COVID-19 Diagnosis Among Adults With and Without Diabetes. *J. Am. Heart Assoc.* **2023**, *12*, e029696. [CrossRef]
81. Domingues, C.P.F.; Rebelo, J.S.; Dionisio, F.; Botelho, A.; Nogueira, T. The Social Distancing Imposed to Contain COVID-19 Can Affect Our Microbiome: A Double-Edged Sword in Human Health. *mSphere* **2020**, *5*, e00716-20. [CrossRef] [PubMed]
82. Bazzoli, A.; Probst, T.M. COVID-19 Moral Disengagement and Prevention Behaviors: The Impact of Perceived Workplace COVID-19 Safety Climate and Employee Job Insecurity. *Saf. Sci.* **2022**, *150*, 105703. [CrossRef] [PubMed]
83. Mulder, C.; Conti, E.; Saccone, S.; Federico, C. Beyond Virology: Environmental Constraints of the First Wave of COVID-19 Cases in Italy. *Env. Sci. Pollut. Res.* **2021**, *28*, 31996–32004. [CrossRef] [PubMed]
84. Birra, D.; Benucci, M.; Landolfi, L.; Merchionda, A.; Loi, G.; Amato, P.; Licata, G.; Quartuccio, L.; Triggiani, M.; Moscato, P. COVID 19: A Clue from Innate Immunity. *Immunol. Res.* **2020**, *68*, 161–168. [CrossRef]
85. Silva, M.J.A.; Rodrigues, Y.C.; Lima, K.V.B.; Lima, L.N.G.C. Innate Immunity to SARS-CoV-2 Infection: A Review. *Epidemiol. Infect.* **2022**, *150*, e142. [CrossRef]
86. Celardo, I.; Pace, L.; Cifaldi, L.; Gaudio, C.; Barnaba, V. The Immune System View of the Coronavirus SARS-CoV-2. *Biol. Direct* **2020**, *15*, 30. [CrossRef]
87. Merx, S.; Neumaier, M.; Wagner, H.; Kirschning, C.J.; Ahmad-Nejad, P. Characterization and Investigation of Single Nucleotide Polymorphisms and a Novel TLR2 Mutation in the Human TLR2 Gene. *Hum. Mol. Genet.* **2007**, *16*, 1225–1232. [CrossRef]
88. Oliveira-Nascimento, L.; Massari, P.; Wetzler, L.M. The Role of TLR2 in Infection and Immunity. *Front. Immunol.* **2012**, *3*, 79. [CrossRef]
89. Van Der Sluis, R.M.; Cham, L.B.; Gris-Oliver, A.; Gammelgaard, K.R.; Pedersen, J.G.; Idorn, M.; Ahmadov, U.; Hernandez, S.S.; Cémalovic, E.; Godsk, S.H.; et al. TLR2 and TLR7 Mediate Distinct Immunopathological and Antiviral Plasmacytoid Dendritic Cell Responses to SARS-CoV-2 Infection. *EMBO J.* **2022**, *41*, e109622. [CrossRef]
90. Khan, S.; Shafiei, M.S.; Longoria, C.; Schoggins, J.; Savani, R.C.; Zaki, H. SARS-CoV-2 Spike Protein Induces Inflammation via TLR2-Dependent Activation of the NF-κB Pathway. *Elife* **2021**, *10*, e68563. [CrossRef]
91. Thuong, N.T.T.; Hawn, T.R.; Thwaites, G.E.; Chau, T.T.H.; Lan, N.T.N.; Quy, H.T.; Hieu, N.T.; Aderem, A.; Hien, T.T.; Farrar, J.J.; et al. A Polymorphism in Human TLR2 Is Associated with Increased Susceptibility to Tuberculous Meningitis. *Genes. Immun.* **2007**, *8*, 422–428. [CrossRef] [PubMed]
92. Texereau, J.; Chiche, J.-D.; Taylor, W.; Choukroun, G.; Comba, B.; Mira, J.-P. The Importance of Toll-Like Receptor 2 Polymorphisms in Severe Infections. *Clin. Infect. Dis.* **2005**, *41*, S408–S415. [CrossRef] [PubMed]
93. Butty, V.; Campbell, C.; Mathis, D.; Benoist, C. DPT-1 Study Group Impact of Diabetes Susceptibility Loci on Progression from Pre-Diabetes to Diabetes in at-Risk Individuals of the Diabetes Prevention Trial-Type 1 (DPT-1). *Diabetes* **2008**, *57*, 2348–2359. [CrossRef] [PubMed]
94. Abrahão Silva, M.J.; Cordeiro Dos Santos, E.; Rodrigues, Y.C.; Batista Lima, K.V.; Costa Lima, L.N.G. Analysis of the Frequency of the Toll-like 2 Gene Polymorphism in Leprosy. *Annu. Res. Rev. Biol.* **2022**, *37*, 1–10. [CrossRef]
95. Qiu, X.; Dong, Y.; Cao, Y.; Luo, Y. Correlation between TLR2, TLR3, TLR4, and TLR9 Polymorphisms and Susceptibility to and Prognosis of Severe Hepatitis among the Newborns. *J. Clin. Lab. Anal.* **2018**, *32*, e22292. [CrossRef]
96. Quan, C.; Ping, J.; Lu, H.; Zhou, G.; Lu, Y. 3DSNP 2.0: Update and Expansion of the Noncoding Genomic Variant Annotation Database. *Nucleic Acids Res.* **2021**, *50*, D950–D955. [CrossRef]
97. Semlali, A.; Parine, N.R.; Al-Numair, N.S.; Almutairi, M.; Hawsawi, Y.M.; Amri, A.A.; Aljebreen, A.M.; Arafah, M.; Almadi, M.A.; Azzam, N.A.; et al. Potential Role of Toll-like Receptor 2 Expression and Polymorphisms in Colon Cancer Susceptibility in the Saudi Arabian Population. *Onco Targets Ther.* **2018**, *11*, 8127–8141. [CrossRef]
98. Manry, J.; Quintana-Murci, L. A Genome-Wide Perspective of Human Diversity and Its Implications in Infectious Disease. *Cold Spring Harb. Perspect. Med.* **2013**, *3*, a012450. [CrossRef]
99. Salamaikina, S.; Karnaushkina, M.; Korchagin, V.; Litvinova, M.; Mironov, K.; Akimkin, V. TLRs Gene Polymorphisms Associated with Pneumonia before and during COVID-19 Pandemic. *Diagnostics* **2022**, *13*, 121. [CrossRef]
100. Doetschman, T. Influence of Genetic Background on Genetically Engineered Mouse Phenotypes. *Methods Mol. Biol.* **2009**, *530*, 423–433. [CrossRef]
101. van der Made, C.I.; Netea, M.G.; van der Veerdonk, F.L.; Hoischen, A. Clinical Implications of Host Genetic Variation and Susceptibility to Severe or Critical COVID-19. *Genome Med.* **2022**, *14*, 96. [CrossRef] [PubMed]
102. El-Sokkary, R.H.; Khater, W.S.; El-Kholy, A.; Eldin, S.M.; Gad, D.M.; Bahgat, S.; Negm, E.E.; El Kholy, J.A.; Mowafy, S.; Mahmoud, E. Compliance of Healthcare Workers to the Proper Use of Personal Protective Equipment during the First Wave of COVID-19 Pandemic. *J. Infect. Public. Health* **2021**, *14*, 1404–1410. [CrossRef] [PubMed]

103. Almeida, C.A.R.P.N.D.; Almeida, G.A.R.P.N.D.; Carvalho, M.R.C.T.D.; Marcolino, A.B.D.L. Aspectos Relacionados à Saúde Mental Dos Profissionais de Saúde Durante a Pandemia Do COVID-19: Uma Revisão Integrativa Da Literatura/Mental Health Aspects of Health Professionals during the Covid-19 Pandemic: An Integrative Literature Review. *Braz. J. Health Rev.* **2020**, *3*, 19481–19491. [CrossRef]
104. Miguel-Puga, J.A.; Cooper-Bribiesca, D.; Avelar-Garnica, F.J.; Sanchez-Hurtado, L.A.; Colin-Martínez, T.; Espinosa-Poblano, E.; Anda-Garay, J.C.; González-Díaz, J.I.; Segura-Santos, O.B.; Vital-Arriaga, L.C.; et al. Burnout, Depersonalization, and Anxiety Contribute to Post-traumatic Stress in Frontline Health Workers at COVID-19 Patient Care, a Follow-up Study. *Brain Behav.* **2021**, *11*, e02007. [CrossRef]

Human Gene 45 (2025) 201426

Contents lists available at ScienceDirect

Human Gene

journal homepage: www.journals.elsevier.com/human-gene

Functional and structural characterization of COVID-19 risk-associated exonic SNPs and identification of novel therapeutic sites: An *in silico* analysis

Marcos Jessé Abrahão Silva [a,*,1], Sebastião Kauã de Sousa Bispo [b,1], Rebecca Lobato Marinho [a], Eliete Costa da Cruz [b], Thiago Pinto Brasil [c], Caroliny Soares Silva [a], Yan Corrêa Rodrigues [a], Cristiane Cunha Frota [d], Diana da Costa Lobato [a], Lilian Cristina Santos Sinfrônio da Silva [b], Everaldina Cordeiro dos Santos [a], Ana Judith Pires Garcia [b], Luana Nepomuceno Gondim Costa Lima [b,*]

[a] Postgraduate Program in Parasitic Biology in the Amazon (PPGBPA), University of Pará State (UEPA), Belém 66087-670, Brazil
[b] Bacteriology and Mycology Section (SABMI), Evandro Chagas Institute (IEC), Ananindeua 67030-000, Brazil
[c] Department of Biomedicine, Federal University of Ceará (UFC), Fortaleza 60441-750, Brazil
[d] Department of Pathology and Legal Medicine, Faculty of Medicine, Federal University of Ceará (UFC), Fortaleza 60441-750, Brazil

ARTICLE INFO

Keywords:
COVID-19
Single nucleotide polymorphism
Computational biology
Genomics

ABSTRACT

The COVID-19 pandemic has highlighted the critical need for effective therapeutic strategies against viral infections, prompting research on the functional characterization of risk-associated single nucleotide polymorphisms (SNPs). This study aimed to analyze exonic SNPs that influence individual susceptibility to COVID-19 through an *in silico* approach. Using a comprehensive methodology, SNPs were retrieved from databases such as Science Direct and PubMed, categorized into intronic, exonic, UTR, splice site, and intergenic types, with a focus on exonic SNPs. Functional analyses were performed using various bioinformatics tools to assess the effects of synonymous and non-synonymous SNPs on mRNA structure, protein stability, protein function, and potential therapeutic sites. The results revealed significant insights into the impact of specific SNPs on COVID-19 susceptibility. For example, the synonymous SNP rs12252 of *IFITM3* was found to have a moderate impact on mRNA structure and binding affinity for microRNAs, while non-synonymous SNPs exhibited varying degrees of functional consequences, with eight variants predicted to be deleterious (with emphasis on the *TYK2* SNP rs34536443 that was predicted to be deleterious in all analyzes). This approach facilitated the identification of novel therapeutic targets. Finally, this research highlights the importance of understanding genetic variations in developing personalized medicine approaches for COVID-19.

1. Introduction

The COVID-19 pandemic has underscored the urgent need for effective therapeutic strategies to combat viral infections (Felsenstein et al., 2020). One promising avenue of research involves the functional and structural characterization of risk-associated exonic single nucleotide polymorphisms (SNPs). Genetic variants known as SNPs are responsible for around 0.1 % of population differences (Caetano, 2009). These SNPs can significantly influence individual susceptibility to COVID-19, making them critical targets for therapeutic intervention. Understanding the implications of these genetic variations is essential for developing personalized medicine approaches that can enhance treatment efficacy and reduce adverse effects (Saiti et al., 2023; Sitinjak et al., 2023; Francès et al., 2024; Yadav et al., 2023).

Recent advancements in *in silico* analysis provide powerful tools for characterizing SNPs associated with COVID-19 risk. By employing computational methods, researchers can analyze vast genomic datasets to identify SNPs that may alter protein function or stability (Hashemi Sheikhshabani et al., 2024; Ariani et al., 2023). This approach not only accelerates the discovery process but also minimizes the costs and time

* Corresponding author.
E-mail addresses: jesseabrahao10@gmail.com (M.J.A. Silva), luanalima@iec.gov.br (L.N.G.C. Lima).
[1] These authors shared first authorship.

https://doi.org/10.1016/j.humgen.2025.201426
Received 17 December 2024; Received in revised form 15 May 2025; Accepted 25 May 2025
Available online 27 May 2025

associated with traditional experimental methods. The integration of bioinformatics tools allows for a comprehensive evaluation of SNPs, facilitating the identification of those with the most significant potential impact on disease progression (Onyango, 2023).

The characterization of exonic SNPs involves assessing their functional consequences on protein structure and activity (DD et al., 2023). Variants that lead to amino acid substitutions can disrupt protein interactions or alter enzymatic functions, thereby influencing cellular responses to viral infections (Słomian et al., 2023; Silva et al., 2023a). Additionally, understanding the structural implications of these SNPs can reveal insights into their role in modulating immune responses, which is particularly relevant in the context of COVID-19 (Delgado-Wicke et al., 2024a).

In addition to identifying risk-associated SNPs, this research aims to uncover novel therapeutic sites that could be targeted for drug development. By mapping the structural landscape of proteins affected by these SNPs, researchers can pinpoint regions that are amenable to therapeutic intervention. This process is crucial for designing drugs that can effectively modulate protein function or inhibit viral replication, ultimately improving patient outcomes (Uslu et al., 2023; Abdelazim et al., 2024).

The significance of this research extends beyond COVID-19, as the methodologies developed can be applied to other infectious diseases and genetic disorders (Fasano et al., 2024; Kumar et al., 2024). Moreover, the integration of machine learning algorithms enhances the predictive power of *in silico* analyses. These algorithms can identify patterns within large datasets, enabling researchers to prioritize SNPs based on their likelihood of contributing to disease risk. This capability is particularly valuable in rapidly evolving fields such as virology, where timely identification of therapeutic targets is essential (Forero et al., 2024). Then, the purpose of the article was to characterize the functional and structural elements of COVID-19 risk-associated exonic SNPs through *in silico* analysis.

2. Material and methods

A workflow for the complete methodology is given in Fig. 1.

2.1. SNP retrieval

Science Direct and PubMed were searched for SNPs linked with COVID-19 until September 2024. Original research publications published in indexed journals served as the source of the data. The bibliographic search criteria for articles included in the databases were: use of MeSH descriptors (COVID-19; Single Nucleotide Polymorphism; SARS-CoV-2) throughout the time period analyzed in the databases and linked to susceptibility and/or severity of COVID-19 in Portuguese, English and/or Spanish. Five categories were created from the retrieved SNPs: intronic, exonic, UTR, splice site and intergenic. For every SNP, details on the mining process were obtained from NCBI dbSNP (accessed on October 01, 2024) (Sayers et al., 2023). These details included global minor allele frequencies (MAF), nucleotide changes, amino acid residual changes for nsSNPs, FASTA sequences, *etc.*

2.2. Characterization of exonic SNPs

Exonic SNPs were divided into synonymous (s) and non-synonymous (n) SNPs. Analyses were performed independently for each of these types. Only missense nsSNPs were considered for their investigation. All SNPs were also evaluated for mutant allele frequency (MAF) at a global level based on the 1000 Genomes Project (Fairley et al., 2020).

2.2.1. Characterization of sSNPs

The COVID-19-associated sSNPs were characterized as below:

2.2.1.1. Prediction of effects on mRNA structure. The prediction of sSNPs effects on mRNA structures were based on the RNAfold server,

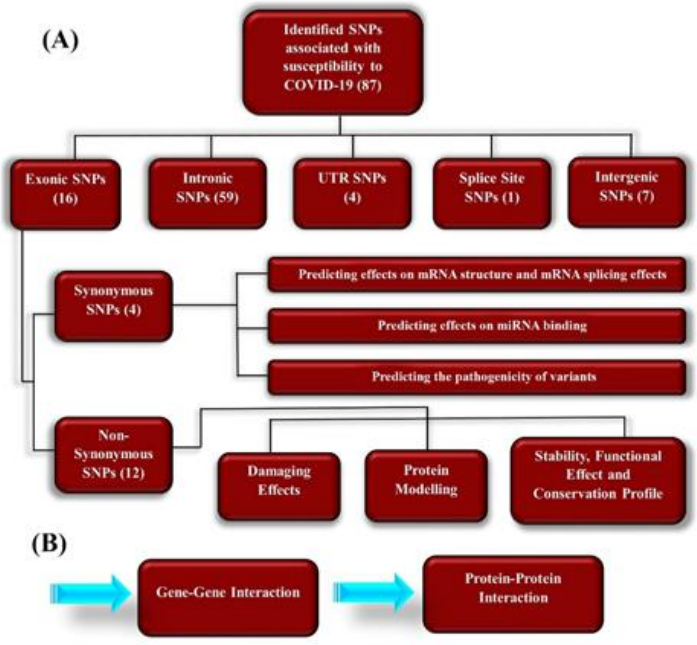


Fig. 1. Methodological flowchart for data analysis.

Table 1
Analysis of mRNA structural, mRNA splicing and miRNA binding effects caused by synonymous SNPs and minor allele frequency across the globe.

Gene	SNP ID	RNAfold (Energy) in ΔΔG (Kcal/mol)	Kinefold (Pseudoknots)	Cyclefold (Noncanonical base pairings)	MaxEntScan in ΔΔG (Kcal/mol)	EX-SKIP (ESS/ESE/ESS:ESSE ratio)	TargetScan (Score)	Global MAF[a]
IFITM3	rs12252	−58.30	−2.6	15.2	−4.33 (equal)	Without mut: 136, 137, 0.99 / with mut: 137, 140, 0.98	−0.37	G = 0.236
TMPRSS2	rs2298659	4.4	−1.9	13.8	−8.42 (equal)	Without mut: 51, 180, 0.28 / with mut: 51, 183, 0.28	−0.39	A = 0.209
IL1RN	rs419598	−1.5	0	16.6	−11.84 (equal)	Without mut: 99, 186, 0.53 / with mut: 96, 182, 0.53	−0.18	C = 0.191
VDR	rs731236	0	0	15.6	−32.88 (equal)	Without: 65, 140, 0.46 / With mut: 65, 143, 0.45	−0.24	G = 0.276

[a] Frequency of existing variant in 1000 Genomes combined population.

Table 2
Analysis of functional effects caused by synonymous SNPs in humans.

Gene	SNP ID	synVEP	TraP	PrDSM	FATHMM-MKL
IFITM3	rs12252	Effect	0.117	0.091	0.064
TMPRSS2	rs2298659	No effect	0.387	0.207	0.028
IL1RN	rs419598	Effect	0.032	0.035	0.038
VDR	rs731236	Effect	0.002	0.019	0.035

comparing sequences with and without the presence of the targeted mutation (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi/; accessed on 01 October 2024). RNAfold is an essential tool in bioinformatics for RNA structure prediction. Algorithm employs partition function calculations in addition to free energy minimization. RNAfold calculates the RNA structure with the lowest free energy, which is considered the most stable configuration of the molecule (Gruber et al., 2008).

Accurate prediction of mRNA structure, including pseudoknots, can help identify pathogenic variants. The presence of pseudoknots can influence the stability and functional activity of mRNA. More complex structures may be required for RNA to perform specific functions, such as regulating translation or interacting with other molecules, such as proteins and microRNAs. The Pseudoknots server was used to predict pseudoknots in mRNA structure (http://rtips.dna.bio.keio.ac.jp/ipknot/; accessed on 01 October 2024). For pseudoknot prediction, it makes use of stochastic folding simulations using folding dynamic algorithms and physical constraint models. It gives the helix tracing graph, folding route video, and lowest free energy structure diagram together with the projected Minimum Free Energy (MFE) (Zhao et al., 2018).

Moreover, non-canonical pairings can contribute to the stability of mRNA structure. They can help stabilize regions that would otherwise be more susceptible to degradation, thereby increasing the longevity of mRNA in the cell. For the prediction of non-canonical base pairings, CycleFold was used (http://rna.urmc.rochester.edu/; accessed on 01 October 2024). It reduces free energy and predicts noncanonical base pairs using nucleotide cyclic patterns. It provides the matrix of pairing probabilities between each nucleotide sequence and the lowest MFE structure (Sloma and Mathews, 2017).

2.2.1.2. Prediction of mRNA splicing effects. MaxEntScan server was used to compare the entropy between sequences with and without the target mutation (http://hollywood.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html; accessed on 01 October 2024). It makes advantage of the maximal entropy principle (MEP)-based Motif approach. Tool delivers a score for the sequence reflecting its strength as splice site (Yeo and Burge, 2004).

SNP Regulatory Elements are regulatory elements of DNA regions that influence gene expression and can be affected by SNPs. The Ex-Skip server was able to predict the possibility of exon silencing (https://ex-skip.img.cas.cz/; accessed on 01 October 2024). The ratio of Exon Splicing Enhancers (ESE)/ Exonic Splicing Silencers (ESSs) found from five distinct models, including RECUE-ESE and FAS-ESSs, is the basis for the predictions. The program estimates the likelihood of exon skipping by comparing the ESE/ESS profiles of a native and mutant sequence (Raponi et al., 2011).

2.2.1.3. Prediction effects on microRNA (miRNA) binding. MiRNAs are small non-coding RNAs that regulate gene expression by binding to complementary sequences in target mRNAs, usually in the 3' UTR (untranslated region) of the mRNA. This interaction can lead to mRNA degradation or inhibition of translation. Synonymous SNPs, although they do not alter the amino acid sequence of the protein, can occur in critical regions that affect mRNA structure, such as miRNA binding sites, and can affect their production and stability. Through sequence alignment prediction, TargetScan is an essential tool in bioinformatics for microRNA target prediction and has been used to predict the influence of variants on gene expression (https://www.targetscan.org/vert_80/; accessed on 01 October 2024). The scores of weighted context++score with a lower negative value indicate a greater prediction of repression (Lewis et al., 2005).

2.2.1.4. Functional effects. PRdsm was used to assess the relationship between sSNPs and disease severity (http://www.xialab.info:8080/PrDSM/; accessed on 02 October 2024). It considers not only the presence of SNPs, but also how these variants can influence gene function and, consequently, the severity of pathological conditions (Cheng et al., 2020). TraP server was also used in this study to assess the functional impact of variants on gene expression and protein translation (https://trap-score.org/home; accessed on 02 October 2024). TraP is especially useful for identifying how synonymous changes can influence translation efficiency or mRNA stability (Gelfman et al., 2017). FATHMM-MKL is a tool that uses hidden Markov models to predict the functional impact of SNPs, and therefore was a method used for sSNPs (https://fathmm.biocompute.org.uk/fathmmMKL.htm/; accessed on 02 October 2024) (Shihab et al., 2015). SynVEP was used to predict the effects of synonymous variants (https://services.bromberglab.org/synvep/home/; accessed on 02 October 2024). It analyzes how these variants can affect mRNA structure, interaction with miRNAs and other regulatory elements (Zeng et al., 2021).

2.2.2. Characterization of nsSNPs

The COVID-19-associated nsSNPs were characterized as below:

2.2.2.1. Most damaging protein. To forecast the impact of nsSNPs on the corresponding proteins, five distinct in-silico methods were employed. Predict SNP 2 (https://loschmidt.chemi.muni.cz/predictsnp2/, accessed on October 1, 2024) (Bendl et al., 2016); Sorting Intolerant from

Tolerant (SIFT) (https://sift.bii.a-star.edu.sg/www/SIFT_seq_submit2.html, accessed on October 1, 2024) (Kumar et al., 2009); SNPs&GO (http://snps.biofold.org/snps-and-go/snps-and-go.html, accessed on October 1, 2024) (Capriotti et al., 2013); Predictor of Human Deleterious SNP (PhD-SNP) (http://snps.biofold.org/phd-snp/phd-snp.html, accessed on October 1, 2024) (Capriotti and Fariselli, 2017); and Polymorphism Phenotyping 2.0 (PolyPhen 2.0) (http://genetics.bwh.harvard.edu/pph2/, accessed on October 1, 2024) (Adzhubei et al., 2013). For every nsSNP, the corresponding proteins' FASTA sequences were supplied, along with the alterations to their amino acid residues. SNPs that had a harmful effect on at least 2 of the 5 servers analyzed went on to the next analyses (Akhtar et al., 2021). Then, nsSNPs were selected such as the results in the software predictions as harmful or deleterious by at least two of the five *in silico* tools were selected for further analysis, which a total of 6 general tools for the validity of this stage of scores for their functional evaluation (Ajith and Subbiah, 2023).

2.2.2.2. Protein stability, structural and functional effects, and conservation profile prediction. To forecast the influence of nsSNPs on the stability of protein, I-Mutant 2.0 (http://folding.biofold.org/i-mutant/i-mutant2.0.html, accessed on 01 October 2024) was used. On a scale of 0 to 10, where 0 and represented the lowest and greatest dependability index, respectively, predictions were given regarding the stability of the mutant protein (Capriotti et al., 2005). The results of the I-Mutant 2 analyses were compared with those of the MuPRO server (https://mupro.proteomics.ics.uci.edu/, accessed on 01 October 2024), regarding protein stability (Cheng et al., 2006). We employed CUPSAT (http://cupsat.tu-bs.de/, accessed on October 2, 2024) to verify the accuracy of the structure-based predictions (Parthiban et al., 2006).

MutPred 1.2 (http://mutpred.mutdb.org/, accessed on 01 October 2024) was used to anticipate the structural and functional impact of nsSNPs on protein. The effect of substituted amino acids on proteins was anticipated, and p values of less than 0.05 and less than 0.01 were regarded as extremely confident and confident, respectively (Pejaver et al., 2020). Using 50 distinct homologous sequences, the ConSurf program (http://consurf.tau.ac.il/2016/, accessed on 01 October 2024) was used to determine the conservation profile of each protein (Ashkenazy et al., 2010).

2.2.2.3. Protein modeling and SNP effect on protein. The SWISS-MODEL server (https://swissmodel.expasy.org/, accessed on 01 October 2024) was used for protein modeling (Waterhouse et al., 2018). Using NCBI BLAST, the templates for every protein that needed to be modeled were found; ultimately, the templates with the highest percentage of identity and coverage were selected from the RCSB Protein Data Bank (http://www.rcsb.org/). Separate command files were created for every protein modeling. Chimera v1.11's built-in functionality was used to model the protein structures of the mutants, and each mutant's unique protein structure's leftover amino acid modifications were done by hand. Following protein modeling, the root-mean-square deviation (RMSD) values for each mutant and wild-type protein were determined using TM-align (https://zhanglab.ccmb.med.umich.edu/TM-align/, accessed on October 5, 2024) (Pettersen et al., 2004).

The impact of gene variations on protein structure was then examined using HOPE (Have (y)Our Protein Explained), which may be found at https://www3.cmbi.umcn.nl/hope/, which was accessed on 01 October 2024. A single mutation point may be analyzed using the website-based program HOPE. The application gathers data from several sources and produces visuals that show the location of mutation (Venselaar et al., 2010).

2.2.2.4. Assessment of relative surface usability. A protein's surface accessibility must be assessed because environmental conditions, such as temperature or chemicals, might alter a protein's ability to fold. The relative surface accessibility of amino acids may be predicted with the aid of NetsurfP 3.0 (https://services.healthtech.dtu.dk/services/NetSurfP-3.0/; accessed on 02 October 2024); results are derived as a Z-score from the network reliability score (Høie et al., 2022).

2.2.3. Gene-gene interactions of associated COVID-19 genes

Using STRING (https://string-db.org/; accessed on 01 October 2024) (Szklarczyk et al., 2022), Pathways Commons (https://www.pathwaycommons.org/; accessed on 02 October 2024) (Rodchenkov et al., 2019), and GeneMANIA (https://genemania.org/; accessed on 02 October 2024) (Mostafavi et al., 2008), the interactions of all the genes chosen for this study were examined. GeneMANIA based its predictions on co-expression, similarity of protein domains, co-localization, and pathways, whereas STRING based its predictions on co-expression, co-occurrence, gene fusion, biochemical, and experimental data. Accessibility to a wide variety of human biochemical processes is possible through Pathway Commons, which groups information from 22 different databases and totals about 4794 pathways and about 2.3 million interactions. To identify the core area genes, a gene list with the official symbols for each gene was submitted. The data was then evaluated.

3. Results

3.1. Qualitative and quantitative characterization of SNPs associated with COVID-19

Fig. 2 shows a pie chart of the distribution and relative proportion of SNPs, according to their type, found to be associated with susceptibility to COVID-19. These included data in the article were presented in spreadsheet form attached in Supplementary Material 1 format, such as their information (Material Suppl. S1).

3.2. Analysis of synonymous SNPs

Among the investigations performed, energy analysis shows the RNA folding energy calculated by RNAfold, which indicates the structural impact of the SNP on the mRNA. More negative values suggest a more stable structure, while positive values indicate a less stable structure. In this sense, the most impacted and most stable mRNA structure was generated by the SNP of the *IFITM3* gene, while the least stable was that of the SNP of the *TMPRSS2* gene, while the *IL1RN* SNP had a small structural change. The Kinefold and CycleFold analyses provide information on the presence of pseudoknots and non-canonical base pairings, respectively. These factors can affect the secondary structure of the

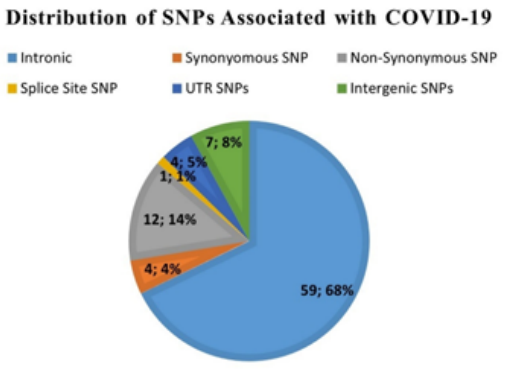


Fig. 2. Distribution of SNPs associated with Poor Prognosis of COVID-19, represented in a pie chart. This figure was generated using Microsoft Excel 2021.

mRNA and influence splicing and translation. Among the values found, the most complex (*IL1RN*) was determined with the absence of pseudoknots and the most positive value of non-canonical bases. Regarding the maximum entropy analysis, when comparing the protein sequences with wild-type and mutant residues, the *VDR* gene and SNP rs731236 had higher negative entropy energy, which may indicate more chances of alternative splicing. Regarding the Ex-Skip and Target Scan servers, higher values indicate a higher probability of exon exclusion and a higher binding affinity to miRNAs, respectively. Regarding the allele frequency of the mutant type of the SNPs analyzed worldwide, the most prevalent was SNP rs731236 of the *VDR* gene. In this context, SNP rs12252 of *IFITM3* was predicted to have a moderate impact on post-transcriptional regulation, while SNP rs419598 of *IL1RN* had a minimal impact. In a general context, the SNP rs12252 in the *IFITM3* gene has the greatest potential impact on mRNA structure, alternative splicing, and miRNA binding, with a relatively high frequency in the global population.

The sSNPs were also analyzed for their functional impact on human protein transcripts. The interpretation of the data suggests that while some SNPs such as rs12252 and rs731236 have potentially significant functional effects, others such as SNP rs2298659 appear to have no relevant impact, while the remaining ones may be predicted to have a moderate impact (See Tables 1 and 2).

3.3. Analysis of non-synonymous SNPs

All non-synonymous SNPs found were characterized as missense type. The functional impact of the analyzed missense SNPs was characterized by each server used in Table 3. Finally, eight nsSNPs (corresponding to eight genes) that were predicted as harmful or deleterious by at least two of the five *in silico* tools were selected for further analysis (Table 3). The Predict SNP 2 tool calculated the parameters of CADD, DANN, FunSeq, GWAVA and FATHMM, which gave it validity as the only consensus score among the others and in a total of 5 tools for the validity of this stage of scores for their functional evaluation.

Table 4 and Table 5 presents the results of the protein stability analysis after the introduction of non-synonymous SNPs, using the CUPSAT tool. The table provides information on the protein stability, the affected secondary structure element, the torsion and the predicted change in Gibbs free energy (ΔΔG).

Swiss-MODEL (https://swissmodel.expasy.org/) was used to visualize the three-dimensional structure of proteins. The produced protein model was created using protein structure templates that were sourced from AlphaFoldDB v4 database. As indicated in Table 6, the outcomes are assessed using the Qualitative Model Energy Analysis (QMEAN). Each modeling result has a very good QMEAN score, ranging from 0.64 to 0.9; a number around 1 indicates the quality of a well-designed model structure. The results demonstrated structural stability in the observed conformations, since there were very low RMSD values. This indicates that the structure of the modeled protein (after the introduction of the SNP) is very similar to the structure of the wild-type protein. RMSD is a measure that quantifies the difference between the positions of the atoms of two structures, and is often used to evaluate the quality of a structural model in relation to the reference structure.

The final models were viewed and studied using Chimera v1.11 (https://www.cgl.ucsf.edu/chimera/, accessed on 16 September 2024). Chimera v1.11 was used for modeling mutant structures by changing the residue of interest. Through TM-align, the RMSD values for each of the mutant proteins were determined for each nsSNP. Fig. 3 displays the modeling findings of the selected pathogenic missense SNPs.

Using the web-based toolset created by Project HOPE, a clearer understanding of how changes in an amino acid residue affect the structure and function of the protein was achieved on Table 7.

The above-mentioned variants of these genes were evaluated for solvent accessibility and stability using the NetsurfP server 2.0. The obtained results are mentioned in Table 8. Mutations in buried sites are

Table 3
Prediction of the functional impact of non-synonymous SNPs using bioinformatics tools.

Gene	SNP ID	Global MAF[a]	Amino Acid Change	SIFT	Poly Phen 2	PhD-SNP	SNPs & GO	Predict SNP 2[#]	FunSeq 2	CADD	DANN	GWAVA	FATHMM
IFIH1	rs1990760	T = 0.356	A946T	Tolerated	Benign	Neutral	Neutral	Neutral	Deleterious	Neutral	Neutral	a	Neutral
NOS3	rs1799983	G-0.823	D298E	Tolerated	Benign	Neutral	Neutral	Neutral	Deleterious	Neutral	Neutral	Deleterious	Neutral
IL-6R	rs2228145	C = 0.293	D358A	Tolerated	Benign	Neutral	Neutral	Neutral	Deleterious	Neutral	Neutral	Neutral	Neutral
TYK2	rs34536443	C = 0.001	P1104A	Deleterious	Pro Damag	Neutral	Neutral	Deleterious	Deleterious	Deleterious	Deleterious	Deleterious	Deleterious
EIFNA4	rs114301457	T = 0.002	F124L	Tolerated	Pro Damag	Disease	Disease	Neutral	Neutral	Deleterious	Neutral	Deleterious	Neutral
PLSCR1	rs343320	A = 0.029	H262Y	Tolerated	Pos Damag	Neutral	Neutral	Deleterious	Deleterious	Deleterious	Deleterious	Deleterious	Deleterious
IFNA10	rs283688148	G = 0.004	W164C	Deleterious	Pro Damag	Disease	Disease	Neutral	Deleterious	Deleterious	Neutral	Deleterious	Neutral
SLC22A31	rs117169628	A = 0.007	P474L	Deleterious	Pro Damag	Disease	Disease	Neutral	Neutral	Neutral	Neutral	a	Deleterious
P2RX7	rs208294	C = 0.530	Y155N	Deleterious	Pro Damag	Disease	Disease	Deleterious	Neutral	Deleterious	Deleterious	a	Deleterious
TLR7	rs179008	T = 0.118	Q11P	Tolerated	Benign	Disease	Disease	Neutral	Neutral	Neutral	Neutral	a	Neutral
IFNL1	rs30461	G = 0.251	N188D	Tolerated	Benign	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral
DDX58	rs10813831	A = 0.184	R7C	Deleterious	Pro Damag	Neutral	Disease	Deleterious	Deleterious	Neutral	Deleterious	Deleterious	Deleterious

[a] Not found. *Frequency of existing variant in 1000 Genomes combined population. [#] Predict SNP 2 was used as only 1 variable that includes other tools to the evaluation.

Table 4
CUPSAT Results of protein stability after nsSNP introduction.

Gene	SNP ID	Stability	SS Element	Torsion	Predicted ΔΔG (kcal/mol)
IFIH1	rs1990760	Stabilising	Helix	Unfavourable	0.47
NOS3	rs1799983	Stabilising	Other (turns, coils, *etc.*)	Unfavourable	0.28
IL-6R	rs2228145	Destabilising	Other (turns, coils, *etc.*)	Favourable	−0.57
TYK2	rs34536443	Stabilising	Helix	Favourable	5.09
EFNA4	rs114301457	Destabilising	Other (turns, coils, *etc.*)	Favourable	−1.27
PLSCR1	rs343320	Destabilising	Sheet	Favourable	−3.16
IFNA10	rs28368148	Destabilising	Helix	Unfavourable	−2.19
SLC22A31	rs117169628	Stabilising	Helix	Favourable	0.88
P2RX7	rs208294	Destabilising	Sheet	Unfavourable	−2.47
TLR7	rs179008	Destabilising	Helix	Unfavourable	−0.87
IFNL1	rs30461	Destabilising	Other (turns, coils, *etc.*)	Favourable	−1.64
DDX58	rs10813831	Destabilising	Helix	Unfavourable	−1.51

Table 5
Predictive analysis of the structural, molecular and evolutionary impact on the protein through the selected nsSNPs.

Gene	SNP ID	I-Mutant 2 (Stability)	MuPRO (Stability)	MutPred		ConSurf Conservation Profile
				PROSITE and ELM Motifs	Molecular Mechanisms	
TYK2	rs34536443	Decrease	Decrease	None	None	Highly conserved, exposed, functional residue
EFNA4	rs114301457	Decrease	Decrease	ELME000133, ELME000249, PS01299	Altered Transmembrane protein	Highly conserved, buried, structural residue
PLSCR1	rs343320	Increase	Decrease	None	None	Highly conserved, exposed, functional residue
IFNA10	rs28368148	Decrease	Decrease	PS00252	Altered Metal binding; Altered Ordered interface; Loss of Allosteric site at W164	Highly conserved, buried, structural residue
SLC22A31	rs117169628	Decrease	Increase	ELME000149, ELME000336, ELME000337, PS00008	Altered Transmembrane protein; Gain of Helix	Highly conserved, exposed, functional residue
P2RX7	rs208294	Decrease	Decrease	None	Altered Transmembrane protein; Altered Metal binding; Loss of Strand; Gain of Disulfide linkage at C152; Loss of Proteolytic cleavage at R151; Gain of GPI-anchor amidation at N158	Average conserved, exposed residue.
TLR7	rs179008	Decrease	Decrease	None	None	Variable, exposed residue.
DDX58	rs10813831	Increase	Decrease	None	None	Variable, exposed residue.

Table 6
Characteristics of query proteins and RMSD value for each one.

Gene	UniProt ID	SNP ID	Amino Acid Change	QMEAN/ GMQE	RMSD (Å)
IFIH1	Q9BYX4	rs1990760	A946T	0.80	0.001
NOS3	P29474	rs1799983	D298E	0.83	0.001
IL-6R	P08887	rs2228145	D358A	0.78	0.000
TYK2	P29597	rs34536443	P1104A	0.75	0.000
EFNA4	P52798	rs114301457	F124L	0.73	0.000
PLSCR1	O15162	rs343320	H262Y	0.70	0.000
IFNA10	P01566	rs28368148	W164C	0.73	0.001
SLC22A31	A6NKX4	rs117169628	P474L	0.62	0.015
P2RX7	Q99572	rs208294	Y155N	0.72	0.001
TLR7	Q9NYK1	rs179008	Q11P	0.78	0.011
IFNL1	Q8IU54	rs30461	N188D	0.68	0.003
DDX58	O95786	rs10813831	R7C	0.78	0.001

more likely to disrupt the protein structure. After further investigation, it was discovered that the mutant type relative solvent accessibility (RSA), and accessible surface area (ASA) values of them have changed compared to the native type.

Fig. 4 shows the protein-protein interactions of the genes of the exonic SNPs analyzed. There were 15 nodes and 27 edges. There was a high interaction score between the nodes of DDX58 with IFIH1 and TLR7; IFNL1 with TLR7, TYK2 and DDX58; IFIH1 with TLR7, IFITM3 and PLSCR1. These proteins were expressed in caspase recruitment (CASP) and C-terminal domains of RIG-I and were divided into two large clusters: C-terminal domain of RIG-I and RIG-I binding; JAK-STAT signaling pathway and interleukin-1 (IL-1) family.

According to the reactome pathways, these proteins had greater activity in decreasing order of strength in the NF-kB activation pathways through the CASP8 and CASP10-mediated FADD/RIP-1 pathway, TRAF3A-dependent IRF activation pathway, TRAF6-mediated IRF7 activation, TRAF6-mediated NF-kB activation, and other cytokine release signaling. Regarding the biological processes involved, these proteins had a greater impact in decreasing order of strength in: regulation of IFN-III formation; virus detection; IFN-III signaling pathway; cytoplasmic pattern recognition receptor signaling pathway in response to the virus; cellular response to dsRNA.

Fig. 5 visually identifies the gene-gene interaction of the genes that had their SNPs associated with COVID-19. They presented a co-expression level of 75.19 %, physical interactions of 10.29 %, protein domain sharing of 9.96 %, co-localization of 4.47 %, and genetic interactions of 0.09 %.

Regarding the interconnection of these genes in the organic pathways, *IFITM3* and *PLSCR1* were associated with the negative regulation pathway of viral genome replication and negative regulation of the viral pathogenicity process. *IFITM3*, *PLSCR1*, and *TMPRSS2* are genes in the

Table 4
CUPSAT Results of protein stability after nsSNP introduction.

Gene	SNP ID	Stability	SS Element	Torsion	Predicted ΔΔG (kcal/mol)
IFIH1	rs1990760	Stabilising	Helix	Unfavourable	0.47
NOS3	rs1799983	Stabilising	Other (turns, coils, *etc.*)	Unfavourable	0.28
IL-6R	rs2228145	Destabilising	Other (turns, coils, *etc.*)	Favourable	−0.57
TYK2	rs34536443	Stabilising	Helix	Favourable	5.09
EFNA4	rs114301457	Destabilising	Other (turns, coils, *etc.*)	Favourable	−1.27
PLSCR1	rs343320	Destabilising	Sheet	Favourable	−3.16
IFNA10	rs28368148	Destabilising	Helix	Unfavourable	−2.19
SLC22A31	rs117169628	Stabilising	Helix	Favourable	0.88
P2RX7	rs208294	Destabilising	Sheet	Unfavourable	−2.47
TLR7	rs179008	Destabilising	Helix	Unfavourable	−0.87
IFNL1	rs30461	Destabilising	Other (turns, coils, *etc.*)	Favourable	−1.64
DDX58	rs10813831	Destabilising	Helix	Unfavourable	−1.51

Table 5
Predictive analysis of the structural, molecular and evolutionary impact on the protein through the selected nsSNPs.

Gene	SNP ID	I-Mutant 2 (Stability)	MuPRO (Stability)	MutPred		ConSurf Conservation Profile
				PROSITE and ELM Motifs	Molecular Mechanisms	
TYK2	rs34536443	Decrease	Decrease	None	None	Highly conserved, exposed, functional residue
EFNA4	rs114301457	Decrease	Decrease	ELME000133, ELME000249, PS01299	Altered Transmembrane protein	Highly conserved, buried, structural residue
PLSCR1	rs343320	Increase	Decrease	None	None	Highly conserved, exposed, functional residue
IFNA10	rs28368148	Decrease	Decrease	PS00252	Altered Metal binding; Altered Ordered interface; Loss of Allosteric site at W164	Highly conserved, buried, structural residue
SLC22A31	rs117169628	Decrease	Increase	ELME000149, ELME000336, ELME000337, PS00008	Altered Transmembrane protein; Gain of Helix	Highly conserved, exposed, functional residue
P2RX7	rs208294	Decrease	Decrease	None	Altered Transmembrane protein; Altered Metal binding; Loss of Strand; Gain of Disulfide linkage at C152; Loss of Proteolytic cleavage at R151; Gain of GPI-anchor amidation at N158	Average conserved, exposed residue.
TLR7	rs179008	Decrease	Decrease	None	None	Variable, exposed residue.
DDX58	rs10813831	Increase	Decrease	None	None	Variable, exposed residue.

Table 6
Characteristics of query proteins and RMSD value for each one.

Gene	UniProt ID	SNP ID	Amino Acid Change	QMEAN/ GMQE	RMSD (Å)
IFIH1	Q9BYX4	rs1990760	A946T	0.80	0.001
NOS3	P29474	rs1799983	D298E	0.83	0.001
IL-6R	P08887	rs2228145	D358A	0.78	0.000
TYK2	P29597	rs34536443	P1104A	0.75	0.000
EFNA4	P52798	rs114301457	F124L	0.73	0.000
PLSCR1	O15162	rs343320	H262Y	0.70	0.000
IFNA10	P01566	rs28368148	W164C	0.73	0.001
SLC22A31	A6NKX4	rs117169628	P474L	0.62	0.015
P2RX7	Q99572	rs208294	Y155N	0.72	0.001
TLR7	Q9NYK1	rs179008	Q11P	0.78	0.011
IFNL1	Q8IU54	rs30461	N188D	0.68	0.003
DDX58	O95786	rs10813831	R7C	0.78	0.001

more likely to disrupt the protein structure. After further investigation, it was discovered that the mutant type relative solvent accessibility (RSA), and accessible surface area (ASA) values of them have changed compared to the native type.

Fig. 4 shows the protein-protein interactions of the genes of the exonic SNPs analyzed. There were 15 nodes and 27 edges. There was a high interaction score between the nodes of DDX58 with IFIH1 and TLR7; IFNL1 with TLR7, TYK2 and DDX58; IFIH1 with TLR7, IFITM3 and PLSCR1. These proteins were expressed in caspase recruitment (CASP) and C-terminal domains of RIG-I and were divided into two large clusters: C-terminal domain of RIG-I and RIG-I binding; JAK-STAT signaling pathway and interleukin-1 (IL-1) family.

According to the reactome pathways, these proteins had greater activity in decreasing order of strength in the NF-kB activation pathways through the CASP8 and CASP10-mediated FADD/RIP-1 pathway, TRAF3A-dependent IRF activation pathway, TRAF6-mediated IRF7 activation, TRAF6-mediated NF-kB activation, and other cytokine release signaling. Regarding the biological processes involved, these proteins had a greater impact in decreasing order of strength in: regulation of IFN-III formation; virus detection; IFN-III signaling pathway; cytoplasmic pattern recognition receptor signaling pathway in response to the virus; cellular response to dsRNA.

Fig. 5 visually identifies the gene-gene interaction of the genes that had their SNPs associated with COVID-19. They presented a co-expression level of 75.19 %, physical interactions of 10.29 %, protein domain sharing of 9.96 %, co-localization of 4.47 %, and genetic interactions of 0.09 %.

Regarding the interconnection of these genes in the organic pathways, *IFITM3* and *PLSCR1* were associated with the negative regulation pathway of viral genome replication and negative regulation of the viral pathogenicity process. *IFITM3*, *PLSCR1*, and *TMPRSS2* are genes in the

Table 7
Identification of The Amino Acid Residue Change Effect of proteins of selected SNPs Using The HOPE Project.

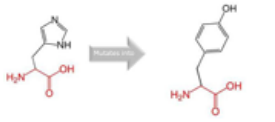
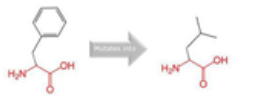
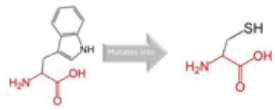
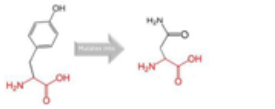
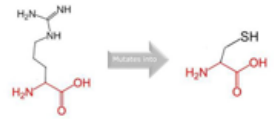
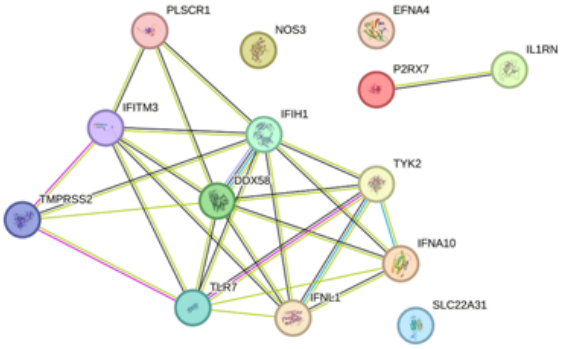
Gene	Protein	Position	Structure	Properties
TYK2	TYK2	P1104A		The mutant residue is smaller, this might lead to loss of interactions. Possibly, this mutation changes a proline with such a function into another residue, thereby disturbing the local structure. The mutation is located within a domain, entitled Protein kinase 2 (JAK).
PLSCR1	PLS1	H262Y		The mutant residue is bigger, this might lead to bumps. The mutation introduces a more hydrophobic residue at this position. This can result in loss of hydrogen bonds and/or disturb correct folding. The mutation is located within a stretch of residues annotated in UniProt as a special motif: Nuclear localization signal, associated with Phospholipid Scramblase Activity.
EFNA4	EFNA4	F124L		The mutant residue is smaller than the wild-type residue. The mutation will cause an empty space in the core of the protein. The mutation is located within domain: Cupredoxin and Ephrin RBD, interfering with Ephrin Receptor Binding.
IFNA10	IFN-10	W164C		The mutant residue is smaller, this might lead to loss of interactions. The mutation is located within the domain: Interferon Alpha/Beta/Delta, associated with Cytokine Receptor Binding and the domain Four-Helical Cytokine-Like, Core.
SLC22A31	S22AV	P474L		The mutant residue is bigger than the wild-type residue. The residue is located on the surface of the protein, mutation of this residue can disturb interactions with other molecules or other parts of the protein. The mutation can affect these interactions and as such affect protein function. It is located at Mfs Transporter Superfamily and Major Facilitator Superfamily Domain, associated with Transmembrane Transporter Activity. The mutation can disturb this special conformation.
P2RX7	P2RX7	Y155N		The mutant residue is smaller than the wild-type residue. This will cause a possible loss of external interactions. The mutation might cause loss of hydrophobic interactions with other molecules on the surface of the protein. It is located at P2X Purinoreceptor domain, with functions: Extracellularly Atp-Gated Monoatomic Cation; Channel Activity; Purinergic Nucleotide Receptor Activity.
TLR7	TLR7	Q11P		The mutant residue is smaller, this might lead to loss of interactions. The mutation introduces a more hydrophobic residue at this position. This can result in loss of hydrogen bonds and/or disturb correct folding. The mutation is located within the signal peptide. Proline disrupts an α-helix when not located at one of the first 3 positions of that helix. In case of the mutation at hand, the helix will be disturbed and this can have severe effects on the structure of the protein. Its domain is unknown.
DDX58	RIG-I	R7C		The mutant residue is smaller than the wild-type residue. The charge of the wild-type residue is lost by this mutation. This can cause loss of interactions with other molecules. The mutant residue is more hydrophobic than the wild-type residue. The mutation is located within domain, annotated as CARD 1 and Death-Like Domain Superfamily.

Table 8
Prediction of relative surface accessibility of these genes through NetsurfP.

Gene	Mutation	NetSurfP			
		Class	RSA	ASA (Å)	Pdisorder
IFIH1	A946T	Exposed	0.49	85	0.09
NOS3	D298E	Exposed	0.53	118	0.02
IL-6R	D358A	Exposed	0.56	73	0.98
TYK2	P1104A	Buried	0.09	12	0.02
EFNA4	F124L	Buried	0.10	21	0.02
PLSCR1	H262Y	Exposed	0.44	115	0.01
IFNA10	W164C	Buried	0.05	8	0.00
SLC22A31	P474L	Buried	0.10	19	0.01
P2RX7	Y155N	Buried	0.25	48	0.03
TLR7	Q11P	Exposed	0.45	72	0.99
IFNL1	N188D	Exposed	0.58	113	0.21
DDX58	R7C	Exposed	0.42	71	0.07

signaling related to COVID-19. Among them, the SNP rs34536443 was associated with greater inflammatory progression of COVID-19 in the Spanish population (Rosa-Baez et al., 2025). The SNP rs114301457 was associated with a higher propensity for COVID-19 in Colombia (Angulo-Aguado et al., 2024). The SNP rs343320 was associated with the critical state of COVID-19 in a UK population (Kousathanas et al., 2021). The SNP rs34536443 was associated in a French cohort with susceptibility to COVID-19 (Kerner and Quintana-Murci, 2022) The SNP rs28368148 in a population sample from Saudi Arabia was associated with a higher risk of COVID-19 (Scaramuzzo et al., 2023; Alghamdi et al., 2021). The SNP rs117169628 was correlated with lower susceptibility and protection to COVID-19 in a Brazilian Amazonian population from Belém-PA (da Costa et al., 2024). SNP rs208294 was correlated with propensity to COVID-19 in a Portuguese population (Lindo et al., 2024). The SNP rs10813831 was associated with COVID-19 severity in an Iranian population (Rokni et al., 2022)O alelo mutante (T) do SNP rs179008 de



Node1	Node2	Node1 accession	Node2 accession	Node 1 annotation	Node 2 annotation	Score
DDX58	IFIH1	ENSP00000369213	ENSP00000497271	Antiviral innate immune response receptor RIG-I	Interferon-induced helicase C domain-containing protein 1	0.96
DDX58	TLR7	ENSP00000369213	ENSP00000370034	Antiviral innate immune response receptor RIG-I	Toll-like receptor 7	0.92
IFNL1	TLR7	ENSP00000329991	ENSP00000370034	Interferon lambda-1	Toll-like receptor 7	0.856
IFIH1	TLR7	ENSP00000497271	ENSP00000370034	Interferon-induced helicase C domain-containing protein 1	Toll-like receptor 7	0.855
IFNL1	TYK2	ENSP00000329991	ENSP00000431885	Interferon lambda-1	Non-receptor tyrosine-protein kinase TYK2	0.852
IFIH1	IFITM3	ENSP00000497271	ENSP00000382707	Interferon-induced helicase C domain-containing protein 1	Interferon-induced transmembrane protein 3	0.757
DDX58	IFNL1	ENSP00000369213	ENSP00000329991	Antiviral innate immune response receptor RIG-I	Interferon lambda-1	0.723
IFIH1	PLSCR1	ENSP00000497271	ENSP00000345494	Interferon-induced helicase C domain-containing protein 1	Phospholipid scramblase 1	0.705

Fig. 4. Protein-protein interaction model of 16 COVID-19-associated genes using STRING. This figure was downloaded as a high-quality image file from STRING v11.0 (https://string-db.org/, accessed on 16 September 2024).

TLR7 foi associado em várias publicações científicas a susceptibilidade à COVID-19 em populações paquistanesa, egípcia e russa (Alseoudy et al., 2022; Khalid et al., 2024; Minashkin et al., 2022).

In addition, the high allele frequency observed in some SNPs, such as rs731236 in *VDR*, suggests that these variants are relevant in broad populations and may have significant public health implications. Identifying common SNPs across different populations may help to better understand the differences in COVID-19 susceptibility and disease severity observed in different ethnic groups. This highlights the need for larger population studies that consider genetic diversity when investigating risk factors for COVID-19 (De Wit et al., 2015).

The integration of advanced bioinformatics tools in the study allowed for a detailed analysis of the functional consequences of SNPs. Predictions about protein stability, evolutionary conservation, and gene interactions provide a comprehensive picture of how these variants may affect cellular functions (Sleator and Walsh, 2010). The use of machine learning algorithms to identify patterns in genomic data also represents a significant advance in genetic research. This approach can accelerate the discovery of new therapeutic targets and facilitate the development of personalized interventions (Terstappen and Reggiani, 2001).

Through mapping the protein structures affected by specific SNPs, researchers can identify regions that are susceptible to modulation by drugs. This strategy is crucial for the development of therapies that can improve clinical outcomes in patients with COVID-19, especially those with genetic variants associated with increased risk. Analysis of gene-gene, protein-protein, and pathway interactions can reveal new therapeutic targets and strategies for clinical interventions. They demonstrate the complexity of the genetic networks involved in the immune response (Gong et al., 2020). Taken together, these insights not only illustrate the diversity of genetic variants associated with COVID-19, but also emphasize the need for a deeper understanding of the functional implications of these SNPs for the development of personalized approaches to treating the disease (Tahery et al., 2021).

Interdisciplinary collaboration between bioinformaticians, clinicians, and pharmacologists is essential to translate *in silico* findings into effective clinical applications. Interaction between these disciplines can facilitate experimental validation of predictions made by computational models and promote the development of new therapeutic approaches based on genetic evidence (Li, 2012). Such collaboration will be essential to address the challenges posed by the pandemic and other infectious diseases in the future. Such interdisciplinary efforts will facilitate the translation of *in silico* findings into clinical applications, ultimately leading to more effective therapies for patients affected by COVID-19 and other viral infections (Lussier et al., 2002).

Furthermore, the findings from the Hope Project highlight the importance of functional characterization of SNPs associated with COVID-19 as a crucial step in identifying new therapeutic targets. The use of *in silico* analyses not only accelerates the discovery process but also provides valuable insights into the interactions between genetics and immune response (Thakur and Shankar, 2016). As we continue to explore the complexities of human genetics in relation to infectious diseases, it is imperative to adopt a holistic approach that considers both genetic and environmental factors in the fight against future pandemics (Ghanavi et al., 2020). The principles of comparative genomics and network-based analysis employed in this study offer a framework for identifying therapeutic targets across various biological contexts. Through leveraging existing genomic data, researchers can uncover new insights into disease mechanisms and potential treatment strategies (Sivashankari and Shanmughavel, 2007).

In silico techniques remain fundamentally imperfect systems, despite recent major improvements in computational tools for analyzing exonic variants. In many situations, predicted disease variants do not reflect actual biological outcomes due to unknown biological complexity or gaps in the amount of reliable and complete genetic datasets (Kumar

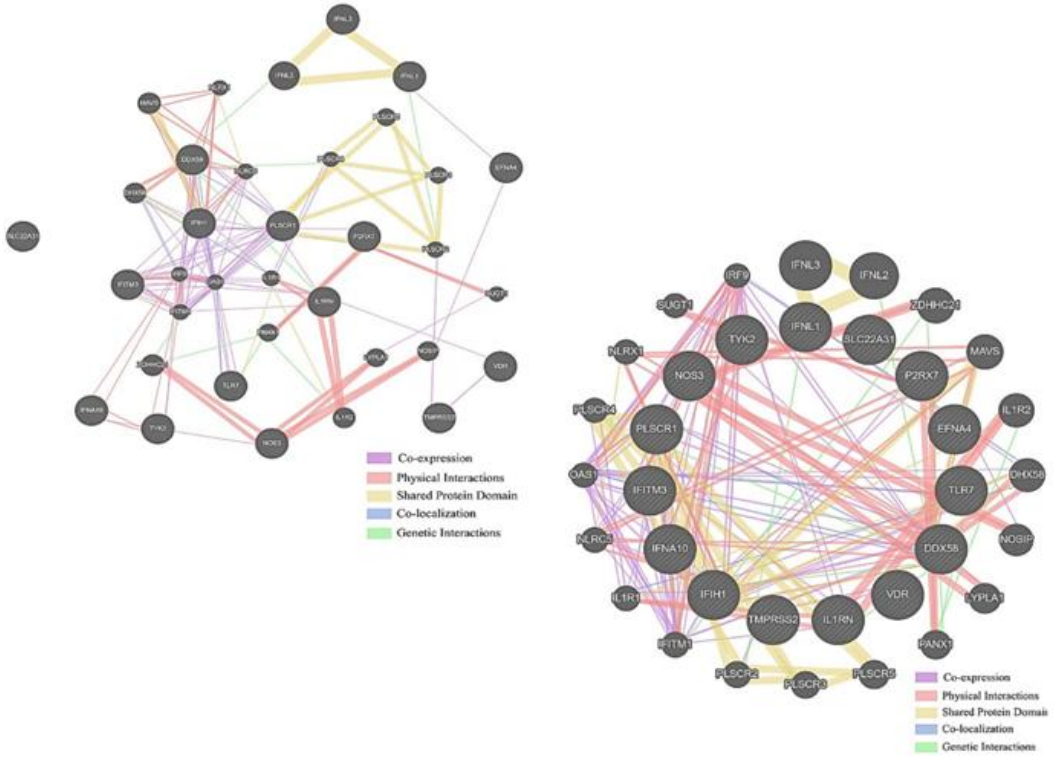


Fig. 5. Gene–gene interaction network generated by GeneMANIA for all the genes interaction types. This figure was downloaded as a high-quality image file from GeneMANIA v3.5.1 (https://genemania.org/, accessed on 16 September 2024).

et al., 2014). It is therefore becoming increasingly crucial that predictions made by *in silico* tools are verified by *in vitro* tests and run by multiple prediction tools using a variety of algorithms and parameters. The need for more sensitive and standardized experimental assays is one reason why there are still very few cases of synonymous variants that have been confirmed empirically. The protein or RNA changes that are found are often modest but substantial (Lin et al., 2023).

Although it, functional experiments provide direct proof of causality beyond statistical association by examining the biological effects of SNPs on gene expression, protein function, or cellular phenotype. Protein binding studies, enzyme activity assessments, electrophoretic mobility shift assays (EMSAs), and reporter gene assays are a few examples. Through their demonstration of their impact in a controlled laboratory environment, these assays aid in validating predicted harmful or regulatory SNPs found by computational techniques. Predicting SNP function and prioritization with a variety of complementary computational methods lowers false positives and boosts confidence in candidate SNPs, validating them (Bora et al., 2023; Majeed et al., 2025; Chorley et al., 2008).

Polygenous Risk Indices (IRP) that combine multiple SNPs can identify individuals at higher risk of severe complications from COVID-19, enabling targeted prevention and intervention strategies. Models that combine SNP data with clinical variables, such as HLA types and comorbidities, have achieved good predictive accuracy (AUC ~0.79), which supports their application in clinical risk assessments. Identification of SNPs linked to immune response genes (such as TYK2 and TLR7) can act as markers for prognosis or response to treatment, aiding in precision therapy and response monitoring. *In silico* screening of SNPs in drug metabolism genes (*e.g.*, CYP3A4/5) can predict individual responses to antivirals such as remdesivir and nirmatrelvir/ritonavir, guiding personalized dosing and minimizing side effects (Alloza-Moral et al., 2025; Delgado-Wicke et al., 2024b; Abdelazim et al., 2025).

5. Conclusion

The investigation of SNPs associated with COVID-19 provides valuable insights into how genetic variations may influence susceptibility and severity of infection. Functional analysis of synonymous and non-synonymous SNPs using advanced bioinformatics tools demonstrated that certain variants, such as sSNPs rs12252 in *IFITM3* and rs731236 in *VDR*, have significant potential to impact mRNA structure and protein function, in addition to the nsSNP rs34536443 of *TYK2* predicted to destabilize the structure but also be capable of deleterious effects on the generated protein. The identification of deleterious SNPs suggests that some variants may predispose individuals to more severe forms of COVID-19, highlighting the need for genetic monitoring in vulnerable populations. These findings underscore the importance of understanding the functional implications of genetic variants in modulating the immune response, which is crucial for the development of personalized therapeutic approaches.

The *in silico* methods employed in this study not only accelerate the discovery of relevant variants but also provide a solid foundation for

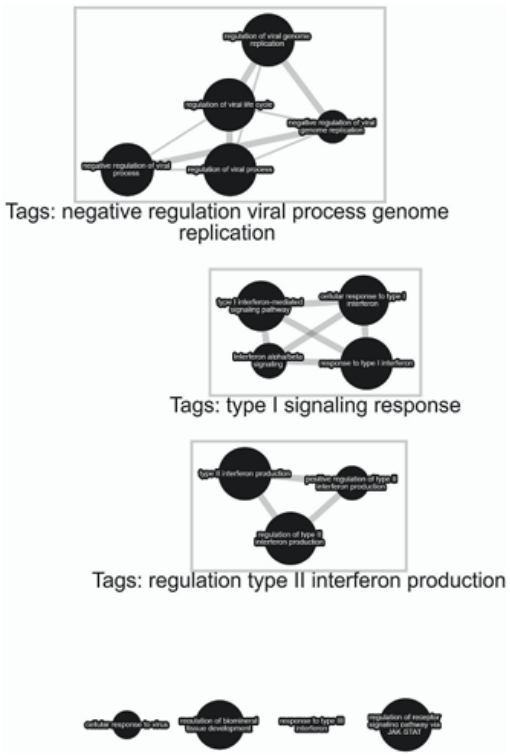


Fig. 6. Pathways connected by same genes. Source: Elaborated through Pathway Commons (2024).

future research in personalized medicine. The ability to predict the functional impact of SNPs through computational analysis may facilitate the identification of novel therapeutic targets and clinical intervention strategies. As we continue to explore the complexities of human genetics in relation to COVID-19 and other infectious diseases, it is critical to foster interdisciplinary collaborations that integrate bioinformatics, molecular biology, and pharmacology. This collaborative approach could lead to the development of more effective and targeted therapies, improving clinical outcomes for patients affected by viral infections.

CRediT authorship contribution statement

Marcos Jessé Abrahão Silva: Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Sebastião Kauã de Sousa Bispo:** Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization. **Rebecca Lobato Marinho:** Methodology, Investigation, Formal analysis, Data curation. **Eliete Costa da Cruz:** Writing – original draft, Resources, Methodology, Investigation, Formal analysis, Data curation. **Thiago Pinto Brasil:** Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis. **Caroliny Soares Silva:** Writing – original draft, Visualization, Validation, Resources, Methodology. **Yan Corrêa Rodrigues:** Validation, Supervision, Investigation, Formal analysis, Data curation. **Cristiane Cunha Frota:** Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Data curation. **Diana da Costa Lobato:** Resources, Investigation, Formal analysis, Data curation. **Lilian Cristina Santos Sinfrônio da Silva:** Writing – original draft, Visualization, Validation, Resources, Methodology. **Everaldina Cordeiro dos Santos:** Writing – review & editing, Visualization, Validation, Software, Resources, Methodology. **Ana Judith Pires Garcia:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Methodology, Investigation, Funding acquisition, Data curation. **Luana Nepomuceno Gondim Costa Lima:** Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Informed consent statement

Not applicable.

Institutional review board statement

Not applicable.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article. The APC will be paid by Evandro Chagas Institute (IEC).

Declaration of competing 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.

Data availability

The original contributions of the study are included in the article. Further inquiries can be directed to the corresponding authors.

References

Abdelazim, A.A., Maged, M., Abdelmaksoud, A.I., Hassanin, S.E., 2024. In-Silico Analysis of nsSNPs in Human CYP3A4/5 and their Effects on Drug-Enzymes Interactions of FDA-Approved COVID-19 Antiviral Drugs: A Pharmacogenetic Study.

Abdelazim, A.A., Maged, M., Abdelmaksoud, A.I., Hassanein, S.E., 2025. In-silico screening and analysis of missense SNPs in human CYP3A4/5 affecting drug-enzyme interactions of FDA-approved COVID-19 antiviral drugs. Sci. Rep. 15, 2153. https://doi.org/10.1038/s41598-025-85595-x.

Adzhubei, I., Jordan, D.M., Sunyaev, S.R., 2013. Predicting functional effect of human missense mutations using PolyPhen-2. Curr. Protoc. Hum. Genet. 76, 7.20. 1-7.20. 41.

Ajith, A., Subbiah, U., 2023. In silico prediction of deleterious non-synonymous SNPs in *STAT3*. Asian Biomed. 17, 185–199. https://doi.org/10.2478/abm-2023-0059.

Akhtar, M., Ali, Y., Islam, Z., Arshad, M., Rauf, M., Ali, M., Maodaa, S.N., Al-Farraj, S.A., El-Serehy, H.A., Jalil, F., 2021. Characterization of rheumatoid arthritis risk-associated SNPs and identification of novel therapeutic sites using an in-silico approach. Biology 10, 501. https://doi.org/10.3390/biology10060501.

Albu-Mohammed, W.H.M., Anvari, E., Fateh, A., 2022. Evaluating the role of BglI Rs739837 and TaqI Rs731236 polymorphisms in vitamin D receptor with SARS-CoV-2 variants mortality rate. Genes 13, 2346. https://doi.org/10.3390/genes13122346.

Alghamdi, J., Alaamery, M., Barhoumi, T., Rashid, M., Alajmi, H., Aljasser, N., Alhendi, Y., Alkhalaf, H., Alqahtani, H., Algablan, O., et al., 2021. Interferon-induced transmembrane protein-3 genetic variant Rs12252 is associated with COVID-19 mortality. Genomics 113, 1733–1741. https://doi.org/10.1016/j.ygeno.2021.04.002.

Alloza-Moral, I., Aldekoa-Etxabe, A., Tulloch-Navarro, R., Fiat-Arriola, A., Mar, C., Urrechaga, E., Ponga, C., Artiga-Folch, I., Garcia-Bediaga, N., Aspichueta, P., et al., 2025. Genetic analysis and predictive modeling of COVID-19 severity in a hospital-based patient cohort. Biomolecules 15, 393. https://doi.org/10.3390/biom15030393.

Alseoudy, M.M., Elgamal, M., Abdelghany, D.A., Borg, A.M., El-Mesery, A., Elzeiny, D., Hammad, M.O., 2022. Prognostic impact of toll-like receptors gene polymorphism on outcome of COVID-19 pneumonia: a case-control study. Clin. Immunol. 235, 108929. https://doi.org/10.1016/j.clim.2022.108929.

Angulo-Aguado, M., Carrillo-Martinez, J.C., Contreras-Bravo, N.C., Morel, A., Parra-Abaunza, K., Usaquén, W., Fonseca-Mendoza, D.J., Ortega-Recalde, O., 2024. Next-generation sequencing of host genetics risk factors associated with COVID-19 severity and long-COVID in Colombian population. Sci. Rep. 14, 8497. https://doi.org/10.1038/s41598-024-57982-3.

Ariani, Y., Muhiardi, I., Shabihah, F., 2023. In silico analysis of single nucleotide polymorphisms related to susceptibility and severity in COVID-19 patients. ejki 271. https://doi.org/10.23886/ejki.11.508.271.

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., Ben-Tal, N., 2010. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res. 38, W529–W533. https://doi.org/10.1093/nar/gkq399.

Bakke, D., Sun, J., 2018. Ancient nuclear receptor VDR with new functions: microbiome and inflammation. Inflamm. Bowel Dis. 24, 1149–1154. https://doi.org/10.1093/ibd/izy092.

Bendl, J., Musil, M., Štourač, J., Zendulka, J., Damborský, J., Brezovský, J., 2016. PredictSNP2: a unified platform for accurately evaluating SNP effects by exploiting the different characteristics of variants in distinct genomic regions. PLoS Comput. Biol. 12, e1004962. https://doi.org/10.1371/journal.pcbi.1004962.

Bora, S., Kukreti, R., Hasija, Y., 2023. An *in-silico* analysis predicting the impact of coding single nucleotide polymorphisms (SNPs) in the human multidrug pump ABCG2. Hum. Gene Ther. 38, 201213. https://doi.org/10.1016/j.humgen.2023.201213.

Caetano, A.R., 2009. SNP markers: basic concepts, applications in animal breeding and management and perspectives for the future. R. Bras. Zootec. 38, 64–71. https://doi.org/10.1590/S1516-35982009001300008.

Capriotti, E., Fariselli, P., 2017. PhD-SNPg: a webserver and lightweight tool for scoring single nucleotide variants. Nucleic Acids Res. 45, W247–W252. https://doi.org/10.1093/nar/gkx369.

Capriotti, E., Fariselli, P., Casadio, R., 2005. I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res. 33, W306–W310.

Capriotti, E., Calabrese, R., Fariselli, P., Martelli, P., Altman, R.B., Casadio, R., 2013. WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation. BMC Genomics 14, S6. https://doi.org/10.1186/1471-2164-14-S3-S6.

Cheng, J., Randall, A., Baldi, P., 2006. Prediction of protein stability changes for single-site mutations using support vector machines. Proteins 62, 1125–1132. https://doi.org/10.1002/prot.20810.

Cheng, N., Li, M., Zhao, L., Zhang, B., Yang, Y., Zheng, C.-H., Xia, J., 2020. Comparison and integration of computational methods for deleterious synonymous mutation prediction. Brief. Bioinform. 21, 970–981. https://doi.org/10.1093/bib/bbz047.

Chorley, B., Wang, X., Campbell, M., Pittman, G., Noureddine, M., Bell, D., 2008. Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies. Mutat. Res. Rev. Mutat. Res. 659, 147–157. https://doi.org/10.1016/j.mrrev.2008.05.001.

da Costa, A.C.A., Albarello Gellen, L.P., Fernandes, M.R., de Coelho, R.C.C., Monte, N., de Moraes, F.C.A., Calderaro, M.C.L., de Freitas, L.M., Matos, J.A., da Fernandes, T.F.S., et al., 2024. Correlation between genomic variants and worldwide COVID-19 epidemiology. J. Pers. Med. 14, 579. https://doi.org/10.3390/jpm14060579.

DD, V.M., Sivaramakrishnan, V., Arvind Kumar, K., 2023. Structural systems biology approach delineate the functional implications of SNPs in exon junction complex interaction network. J. Biomol. Struct. Dyn. 41, 11969–11986. https://doi.org/10.1080/07391102.2022.2164355.

De Wit, P., Pespeni, M.H., Palumbi, S.R., 2015. SNP genotyping and population genomics from expressed sequences – current advances and future possibilities. Mol. Ecol. 24, 2310–2323. https://doi.org/10.1111/mec.13165.

Delgado-Wicke, P., Fernández De Córdoba-Oñate, S., Roy-Vallejo, E., Alegría-Carrasco, E., Rodríguez-Serrano, D.A., Lamana, A., Montes, N., Nicolao-Gómez, A., Carracedo-Rodríguez, R., Marcos-Jiménez, A., et al., 2024a. Genetic variants regulating the immune response improve the prediction of COVID-19 severity provided by clinical variables. Sci. Rep. 14, 20728. https://doi.org/10.1038/s41598-024-71476-2.

Delgado-Wicke, P., Fernández De Córdoba-Oñate, S., Roy-Vallejo, E., Alegría-Carrasco, E., Rodríguez-Serrano, D.A., Lamana, A., Montes, N., Nicolao-Gómez, A., Carracedo-Rodríguez, R., Marcos-Jiménez, A., et al., 2024b. Genetic variants regulating the immune response improve the prediction of COVID-19 severity provided by clinical variables. Sci. Rep. 14, 20728. https://doi.org/10.1038/s41598-024-71476-2.

Fairley, S., Lowy-Gallego, E., Perry, E., Flicek, P., 2020. The international genome sample resource (IGSR) collection of open human genomic variation resources. Nucleic Acids Res. 48, D941–D947. https://doi.org/10.1093/nar/gkz836.

Fasano, C., Lepore Signorile, M., De Marco, K., Forte, G., Disciglio, V., Sanese, P., Grossi, V., Simone, C., 2024. In silico deciphering of the potential impact of variants of uncertain significance in hereditary colorectal cancer syndromes. Cells 13, 1314. https://doi.org/10.3390/cells13161314.

Felsenstein, S., Herbert, J.A., McNamara, P.S., Hedrich, C.M., 2020. COVID-19: immunology and treatment options. Clin. Immunol. 215, 108448. https://doi.org/10.1016/j.clim.2020.108448.

Forero, D.A., Bonilla, D.A., González-Giraldo, Y., Patrinos, G.P., 2024. An overview of key online resources for human genomics: a powerful and open toolbox for *in silico* research. Brief. Funct. Genomics, elae029. https://doi.org/10.1093/bfgp/elae029.

Frances, M.P., Vila-Vecilla, L., Russo, V., Caetano Polonini, H., De Souza, G.T., 2024. Utilising SNP association analysis as a prospective approach for personalising androgenetic alopecia treatment. Dermatol. Ther. (Heidelb) 14, 971–981. https://doi.org/10.1007/s13555-024-01142-y.

Gelfman, S., Wang, Q., McSweeney, K.M., Ren, Z., La Carpia, F., Halvorsen, M., Schoch, K., Ratzon, F., Heinzen, E.L., Boland, M.J., et al., 2017. Annotating pathogenic non-coding variants in genic regions. Nat. Commun. 8, 236. https://doi.org/10.1038/s41467-017-00141-2.

Ghanavi, J., Farnia, P., Farnia, P., Velayati, A., 2020. Human genetic background in susceptibility to tuberculosis. Int. J. Mycobacteriol. 9, 239. https://doi.org/10.4103/ijmy.ijmy_118_20.

Ginsburg, G.S., Willard, H.F., 2009. Genomic and personalized medicine: foundations and applications. Transl. Res. 154, 277–287. https://doi.org/10.1016/j.trsl.2009.09.005.

Gong, F.-C., Ji, R., Wang, Y.-M., Yang, Z.-T., Chen, Y., Mao, E.-Q., Chen, E.-Z., 2020. Identification of potential biomarkers and immune features of sepsis using bioinformatics analysis. Mediat. Inflamm. 2020, 1–12. https://doi.org/10.1155/2020/3432587.

Greiller, C.L., Martineau, A.R., 2015. Modulation of the immune response to respiratory viruses by vitamin D. Nutrients 7, 4240–4270. https://doi.org/10.3390/nu7064240.

Gruber, A.R., Lorenz, R., Bernhart, S.H., Neubock, R., Hofacker, I.L., 2008. The Vienna RNA Websuite. Nucleic Acids Res. 36, W70–W74. https://doi.org/10.1093/nar/gkn188.

Hashemi Sheikhshabani, S., Ghafouri-Fard, S., Amini-Farsani, Z., Modarres, P., Khazaei Feyzabad, S., Amini-Farsani, Z., Shaygan, N., Omrani, M.D., 2024. In silico prediction of functional SNPs interrupting antioxidant defense genes in relation to COVID-19 progression. Biochem. Genet. https://doi.org/10.1007/s10528-024-10705-9.

Høie, M.H., Kiehl, E.N., Petersen, B., Nielsen, M., Winther, O., Nielsen, H., Hallgren, J., Marcatili, P., 2022. NetSurfP-3.0: accurate and fast prediction of protein structural features by protein language models and deep learning. Nucleic Acids Res. 50, W510–W515. https://doi.org/10.1093/nar/gkac439.

Kerner, G., Quintana-Murci, L., 2022. The genetic and evolutionary determinants of COVID-19 susceptibility. Eur. J. Hum. Genet. 30, 915–921. https://doi.org/10.1038/s41431-022-01141-7.

Khalid, B., Farukh, S., Kumar, A., Baig, S., Shahid, M.A., 2024. Synonymous variant of TLR7 at restriction site Rs864058 identified in Covid 19 Pakistani patients. Am. J. Blood Res. 14, 6–13. https://doi.org/10.62347/YSKN6673.

Kousathanas, A., Pairo-Castineira, E., Rawlik, K., Stuckey, A., Odhams, C.A., Walker, S., Russell, C.D., Malinauskas, T., Millar, J., Elliott, K.S., et al., 2021. Whole Genome Sequencing Identifies Multiple Loci for Critical Illness Caused by COVID-19.

Kumar, P., Henikoff, S., Ng, P.C., 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. 4, 1073–1081. https://doi.org/10.1038/nprot.2009.86.

Kumar, A., Rajendran, V., Sethumadhavan, R., Shukla, P., Tiwari, S., Purohit, R., 2014. Computational SNP analysis: current approaches and future prospects. Cell Biochem. Biophys. 68, 233–239. https://doi.org/10.1007/s12013-013-9705-6.

Kumar, R., Jayaraman, M., Ramadas, K., Chandrasekaran, A., 2024. Computational identification and analysis of deleterious non-synonymous single nucleotide polymorphisms (nsSNPs) in the human POR gene: a structural and functional impact. J. Biomol. Struct. Dyn. 42, 1518–1532. https://doi.org/10.1080/07391102.2023.2211674.

Lewis, B.P., Burge, C.B., Bartel, D.P., 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are MicroRNA targets. Cell 120, 15–20. https://doi.org/10.1016/j.cell.2004.12.035.

Li, B., 2012. Clinical Genomic Analysis and Diagnosis –Genomic Analysis Ex Vivo, in Vitro and in Silico, 2. CMD, pp. 37–44. https://doi.org/10.5923/j.cmd.20120204.04.

Lin, B.C., Katneni, U., Jankowska, K.I., Meyer, D., Kimchi-Sarfaty, C., 2023. In silico methods for predicting functional synonymous variants. Genome Biol. 24, 126. https://doi.org/10.1186/s13059-023-02966-1.

Lindo, J., Nogueira, C., Soares, R., Cunha, N., Almeida, M.R., Rodrigues, L., Coelho, P., Rodrigues, F., Cunha, R.A., Gonçalves, T., 2024. Genetic polymorphisms of P2RX7 but not of ADORA2A are associated with the severity of SARS-CoV-2 infection. Int. J. Mol. Sci. 25, 6135. https://doi.org/10.3390/ijms25116135.

Lussier, Y.A., Sarkar, I.N., Cantor, M., 2002. An integrative model for in-silico clinical-genomics discovery science. Proc. AMIA Symp. 469–473.

Majeed, S., Moin, H., Waseem, M., Khalid, Z., Wajid Abbasi, S., Rasool, K., 2025. Identification of candidate nsSNPs of the human FNDC5 gene and their structural and functional consequences using in silico analysis. Sci. Rep. 15, 7681. https://doi.org/10.1038/s41598-024-83254-1.

Minashkin, M.M., Grigortsevich, N.Y., Kamaeva, A.S., Barzanova, V.V., Traspov, A.A., Godkov, M.A., Ageev, F.A., Petrikov, S.S., Pozdnyakova, N.V., 2022. The role of genetic factors in the development of acute respiratory viral infection COVID-19: predicting severe course and outcomes. Biomedicines 10, 549. https://doi.org/10.3390/biomedicines10030549.

Mostafavi, S., Ray, D., Warde-Farley, D., Grouios, C., Morris, Q., 2008. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. Genome Biol. 9, S4. https://doi.org/10.1186/gb-2008-9-s1-s4.

Munshi, R., Hussein, M.H., Toraih, E.A., Elshazli, R.M., Jardak, C., Sultana, N., Youssef, M.R., Omar, M., Attia, A.S., Fawzy, M.S., et al., 2021. Vitamin D insufficiency as a potential culprit in critical COVID-19 patients. J. Med. Virol. 93, 733–740. https://doi.org/10.1002/jmv.26360.

Onyango, O.H., 2023. In silico models for anti-COVID-19 drug discovery: a systematic review. Adv. Pharmacol. Pharm. Sci. 2023, 1–15. https://doi.org/10.1155/2023/4562974.

Parthiban, V., Gromiha, M.M., Schomburg, D., 2006. CUPSAT: prediction of protein stability upon point mutations. Nucleic Acids Res. 34, W239–W242. https://doi.org/10.1093/nar/gkl190.

Pejaver, V., Urresti, J., Lugo-Martinez, J., Pagel, K.A., Lin, G.N., Nam, H.-J., Mort, M., Cooper, D.N., Sebat, J., Iakoucheva, L.M., et al., 2020. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nat. Commun. 11, 5918. https://doi.org/10.1038/s41467-020-19669-x.

Pereira, M., Dantas Damascena, A., Galvão Azevedo, L.M., De Almeida Oliveira, T., Da Mota Santana, J., 2022. Vitamin D deficiency aggravates COVID-19: systematic review and meta-analysis. Crit. Rev. Food Sci. Nutr. 62, 1308–1316. https://doi.org/10.1080/10408398.2020.1841090.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612. https://doi.org/10.1002/jcc.20084.

Raponi, M., Kralovicova, J., Copson, E., Divina, P., Eccles, D., Johnson, P., Baralle, D., Vorechovsky, I., 2011. Prediction of single-nucleotide substitutions that result in exon skipping: identification of a splicing silencer in *BRCA1* exon 6. Hum. Mutat. 32, 436–444. https://doi.org/10.1002/humu.21458.

Rodchenkov, I., Babur, O., Luna, A., Aksoy, B.A., Wong, J.V., Fong, D., Franz, M., Siper, M.C., Cheung, M., Wrana, M., et al., 2019. Pathway commons 2019 update: integration, analysis and exploration of pathway data. Nucleic Acids Res., gkz946 https://doi.org/10.1093/nar/gkz946.

Rokni, M., Sarhadi, M., Heidari Nia, M., Mohamed Khosroshahi, L., Asghari, S., Sargazi, S., Mirinejad, S., Saravani, R., 2022. Single nucleotide polymorphisms located in TNFA, IL1RN, IL6R, and IL6 genes are associated with COVID-19 risk and severity in an Iranian population. Cell Biol. Int. 46, 1109–1127. https://doi.org/10.1002/cbin.11807.

Rosa-Baez, C., Borrego-Yaniz, G., Rodriguez-Martin, I., Kerick, M., Acosta-Herrera, M., Martín, J., Ortiz-Fernández, L., 2025. Cross-trait GWAS in COVID-19 and systemic sclerosis reveals novel genes implicated in fibrotic and inflammation pathways. Rheumatology, keaf028. https://doi.org/10.1093/rheumatology/keaf028.

Rusiñol, L., Puig, L., 2023. Tyk2 targeting in immune-mediated inflammatory diseases. Int. J. Mol. Sci. 24, 3391. https://doi.org/10.3390/ijms24043391.

Saiti, A., Giannopoulos-Dimitriou, A., Kazakos, I., Galatou, E., Vizirianakis, I.S., 2023. Systems pharmacology and network analysis to advance pharmacogenomics and precision medicine decisions in type-2 diabetes therapy. Futur. Pharmacol. 3, 329–363. https://doi.org/10.3390/futurepharmacol3010021.

Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, D.C., Farrell, C.M., Feldgarden, M., Fine, A.M., Funk, K., 2023. Database resources of the National Center for biotechnology information in 2023. Nucleic Acids Res. 51, D29.

Scaramuzzo, G., Nucera, F., Asmundo, A., Messina, R., Mari, M., Montanaro, F., Johansen, M.D., Monaco, F., Fadda, G., Tuccari, G., et al., 2023. Cellular and molecular features of COVID-19 associated ARDS: therapeutic relevance. J. Inflamm. (Lond) 20, 11. https://doi.org/10.1186/s12950-023-00333-2.

Shi, G., Kenney, A.D., Kudryashova, E., Zani, A., Zhang, L., Lai, K.K., Hall-Stoodley, L., Robinson, R.T., Kudryashov, D.S., Compton, A.A., et al., 2021. Opposing activities of IFITM proteins in SARS-CoV-2 infection. EMBO J. 40, e106501. https://doi.org/10.15252/embj.2020106501.

Shihab, H.A., Rogers, M.F., Gough, J., Mort, M., Cooper, D.N., Day, I.N.M., Gaunt, T.R., Campbell, C., 2015. An integrative approach to predicting the functional effects of non-coding and coding sequence variation. Bioinformatics 31, 1536–1543. https://doi.org/10.1093/bioinformatics/btv009.

Silva, M.J.A., Silva, C.S., Da Silva Vieira, M.C., Dos Santos, P.A.S., Frota, C.C., Lima, K.V. B., Lima, L.N.G.C., 2023a. The relationship between TLR3 Rs3775291 polymorphism and infectious diseases: a meta-analysis of case-control studies. Genes 14, 1311. https://doi.org/10.3390/genes14071311.

Silva, M.J.A., Ribeiro, L.R., Gouveia, M.I.M., dos Marcelino, B.R., dos Santos, C.S., Lima, K.V.B., Lima, L.N.G.C., 2023b. Hyperinflammatory response in COVID-19: a systematic review. Viruses 15, 553. https://doi.org/10.3390/v15020553.

Sitinjak, B.D.P., Murdaya, N., Rachman, T.A., Zakiyah, N., Barliana, M.I., 2023. The potential of single nucleotide polymorphisms (SNPs) as biomarkers and their association with the increased risk of coronary heart disease: a systematic review. VHRM 19, 289–301. https://doi.org/10.2147/VHRM.S405039.

Sivashankari, S., Shanmughavel, P., 2007. Comparative genomics - a perspective. Bioinformation 1, 376. https://doi.org/10.6026/97320630001376.

Sleator, R.D., Walsh, P., 2010. An overview of in silico protein function prediction. Arch. Microbiol. 192, 151–155. https://doi.org/10.1007/s00203-010-0549-9.

Sloma, M.F., Mathews, D.H., 2017. Base pair probability estimates improve the prediction accuracy of RNA non-canonical base pairs. PLoS Comput. Biol. 13, e1005827. https://doi.org/10.1371/journal.pcbi.1005827.

Słomian, D., Szyda, J., Dobosz, P., Stojak, J., Michalska-Foryszewska, A., Sypniewski, M., Liu, J., Kotlarz, K., Suchocki, T., Mroczek, M., et al., 2023. Better safe than sorry—whole-genome sequencing indicates that missense variants are significant in susceptibility to COVID-19. PLoS One 18, e0279356. https://doi.org/10.1371/journal.pone.0279356.

Smith, A.J.P., Humphries, S.E., 2009. Cytokine and cytokine receptor gene polymorphisms and their functionality. Cytokine Growth Factor Rev. 20, 43–59. https://doi.org/10.1016/j.cytogfr.2008.11.006.

Sunyaev, S., Hanke, J., Aydin, A., Wirkner, U., Zastrow, I., Reich, J., Bork, P., 1999. Prediction of nonsynonymous single nucleotide polymorphisms in human disease-associated genes. J. Mol. Med. 77, 754–760. https://doi.org/10.1007/s001099900059.

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N.T., Pyysalo, S., et al., 2022. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 51, D638–D646. https://doi.org/10.1093/nar/gkac1000.

Tahery, N., Khodadost, M., Jahani Sherafat, S., Rezaei Tavirani, M., Ahmadi, N., Montazer, F., Rezaei Tavirani, M., Naderi, N., 2021. C-reactive protein as a possible marker for severity and mortality of COVID-19 infection. Gastroenterol. Hepatol. Bed Bench 14, S118–S122.

Terstappen, G.C., Reggiani, A., 2001. In silico research in drug discovery. Trends Pharmacol. Sci. 22, 23–26. https://doi.org/10.1016/S0165-6147(00)01584-4.

Thakur, R., Shankar, J., 2016. In silico analysis revealed high-risk single nucleotide polymorphisms in human pentraxin-3 gene and their impact on innate immune response against microbial pathogens. Front. Microbiol. 7. https://doi.org/10.3389/fmicb.2016.00192.

Uslu, K., Ozcelik, F., Zararsiz, G., Eldem, V., Cephe, A., Sahin, I.O., Yuksel, R.C., Sipahioglu, H., Ozer Simsek, Z., Baspinar, O., et al., 2023. Deciphering the host genetic factors conferring susceptibility to severe COVID-19 using exome sequencing. Genes Immun. 25, 14–42. https://doi.org/10.1038/s41435-023-00232-9.

Venselaar, H., Te Beek, T.A., Kuipers, R.K., Hekkelman, M.L., Vriend, G., 2010. Protein structure analysis of mutations causing inheritable diseases. An e-science approach with life scientist friendly interfaces. BMC Bioinformatics 11, 548. https://doi.org/10.1186/1471-2105-11-548.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T.A.P., Rempfer, C., Bordoli, L., et al., 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296–W303. https://doi.org/10.1093/nar/gky427.

Xu, F., Wang, G., Zhao, F., Huang, Y., Fan, Z., Mei, S., Xie, Y., Wei, L., Hu, Y., Wang, C., et al., 2022. IFITM3 inhibits SARS-CoV-2 infection and is associated with COVID-19 susceptibility. Viruses 14, 2553. https://doi.org/10.3390/v14112553.

Yadav, A., Srivastava, S., Tyagi, S., Krishna, N., Katara, P., 2023. In-silico mining to glean SNPs of pharmaco-clinical importance: an investigation with reference to the Indian populated SNPs. In Silico Pharmacol. 11, 17. https://doi.org/10.1007/s40203-023-00154-4.

Yeo, G., Burge, C.B., 2004. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J. Comput. Biol. 11, 377–394. https://doi.org/10.1089/1066527041410418.

Zabihi Rizi, F., Ghorbani, A., Zahtab, P., Darbaghshahi, N.N., Ataee, N., Pourhamzeh, P., Hamzei, B., Dolatabadi, N.F., Zamani, A., Hooshmand, M., 2023. TYK2 single-nucleotide variants associated with the severity of COVID-19 disease. Arch. Virol. 168, 119. https://doi.org/10.1007/s00705-023-05729-2.

Zeng, Z., Aptekmann, A.A., Bromberg, Y., 2021. Decoding the effects of synonymous variants. Nucleic Acids Res. 49, 12673–12691. https://doi.org/10.1093/nar/gkab1159.

Zhao, Y., Wang, J., Zeng, C., Xiao, Y., 2018. Evaluation of RNA secondary structure prediction for both base-pairing and topology. Biophys. Rep. 4, 123–132. https://doi.org/10.1007/s41048-018-0058-y.

APÊNDICE 4

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

(Termo apresentado conforme normas da Resolução 466/2012)

Título da pesquisa: "*Análise da resposta ao SARS-COV-2 em relação aos achados radiológicos e/ou à susceptibilidade genética individual*"

Prezado(a) Senhor(a):

Gostaríamos de convidá-lo (a) a participar, como voluntário, da pesquisa que tem como título **"*Análise da resposta ao SARS-COV-2 em relação aos achados radiológicos e/ou à susceptibilidade genética individual.*"** a ser realizada em centros de saude de Belém e Ananindeua em parceria com o Instituto Evandro Chagas. O objetivo principal da pesquisa é saber se você tem algum fator genético no sangue que podem tornar mais fácil de você adquirir a COVID19 e agravar. Também, no caso de você ter adoecido e feito tomografias de tórax (TC), será avaliado os principais achados radiológicos da sua TC compatíveis com a COVID- 19 em relação à sua genética, sintomatologia, evolução e desfecho clínico da COVID-19.

É muito importante fazer essa pesquisa porque não temos informações sobre se as pessoas que não adoecem são geneticamente diferentes das pessoas que adoecem e se a genética interferiu nos achados radiológicos da TC. Isso pode ajudar as autoridades de saúde a conhecer melhor as características clínicas da doença e o motivo pelo qual as pessoas estão adoecendo. Esta pesquisa será realizada com profissionais que trabalharam em centros de saúde de Belém e Ananindeua por pelo menos trinta dias do período compreendido entre 10/03 a 10/07/2020 que tiveram ou não a COVID19. Serão usadas nesta pesquisa amostras de sangue que serão coletadas na Policlínica Metropolitana de Belém. As amostras de sangue serão coletadas do braço com seringa e agulha descartáveis para analisarmos o gene do Receptor de vitamina D (VDR), genes de citocinas, inflamossomo e, Toll-like Receptores (TLR), enzima conversora de angiotensina 2 (ACE-2) e a serina protease

transmembranar 2 (TMPRSS2). Esses genes serão analisados para saber se tem um fator genético que pode tornar mais fácil a pessoa adoecer. Serão feitas algumas perguntas para você responder e essas perguntas ficaram registradas em um questionário.

A sua participação é muito importante e, em respeito a você, lhe damos as seguintes explicações:

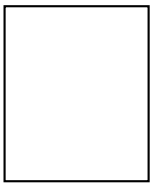
- Você não é obrigado a participar da pesquisa e pode retirar seu consentimento a qualquer momento, e caso não aceite, você não será maltratado ou prejudicado.
- Como resultado dessa pesquisa saberemos se você tem um fator genético que pode o tornar mais susceptível ao adoecimento. Os resultados do fator genético, você poderá ter acesso se desejar, pois serão utilizados apenas para a pesquisa.
- Esta pesquisa não trará um benefício direto para você, mas trará benefícios indiretos como: os fatores genéticos que podem favorecer determinados achados radiológicos da TC e o adoecimento, o que é muito importante para as atividades de controle da doença pelas secretarias de saúde.
- Os riscos desta pesquisa referem-se à coleta de sangue pela picada da agulha que produzirá um desconforto suportável e de curta duração. Poderá ocorrer inchaço no local da picada, durante ou após a coleta. Também poderá entrar um germe no local da coleta, no entanto serão tomados todos os cuidados de limpeza recomendados pelo Ministério da Saúde. Outro risco está relacionado ao sigilo de seus dados, mas garantimos que suas informações serão utilizadas somente para os fins desta pesquisa e serão tratadas com o mais absoluto sigilo e confidencialidade, de modo a preservar a sua identidade.
- A explicação da pesquisa, e aplicação do Termo de Consentimento serão feitas num lugar onde você possa se sentir à vontade para tirar suas dúvidas.
- Esta pesquisa não gera nenhuma despesa aos participantes. Para participar deste estudo você não terá nenhum custo, nem receberá qualquer vantagem financeira. Sua participação é voluntária.
- Você terá garantia de reparação de dano, comprovadamente decorrente da participação na pesquisa, conforme previsto em lei.

Caso o(a) senhor(a) tenha dúvidas ou necessite de maiores esclarecimentos pode nos contactar: **Bruna Venturieri –** Pesquisadora Responsável pelo projeto, encontrada na **Seção Policlínica Metropolitana de Belém, Tv. Perebebuí, 2623 - Marco, Belém - PA; Telefone: (91)991492702**, E-mail: venturierib@icloud.com, ou procurar o **Comitê de Ética em Pesquisa Envolvendo Seres Humanos da Policlínica Metropolitana de Belém**, localizado no Prédio da Policlínica Metropolitana de Belém, Tv. Perebebuí, 2623 - Marco, Belém - PA, no Telefone: (91) 40050510. Horário de funcionamento: segunda a sexta-feira, das 8hrs ÀS 17hrs.

Este termo deverá ser preenchido em duas vias de igual teor, sendo uma delas, devidamente preenchida, assinada e entregue ao(a) senhor(a).

CONSENTIMENTO DA PESSOA COMO PARTICIPANTE DA PESQUISA

Eu,_____,
RG/CPF_____ concordo em participar do estudo "***Análise da resposta ao SARS-COV-2 em relação aos achados radiológicos e/ou à susceptibilidade genética individual***", sob a responsabilidade da Pesquisadora Bruna Venturieri, como participante voluntário. Fui devidamente informado e esclarecido sobre a pesquisa, os procedimentos nela envolvidos, assim como os possíveis riscos e benefícios decorrentes de minha participação. Foi me garantido que posso retirar meu consentimento a qualquer momento, sem que isto leve à qualquer penalidade ou interrupção de meu acompanhamento/ assistência/ tratamento.
Local e data: _
Assinatura do participante ou responsável: _

Assinatura Dactiloscópica 

Assinatura do Pesquisador Responsável: _

Presenciamos a solicitação de consentimento, esclarecimento sobre a pesquisa e aceite do sujeito em participar.

Testemunhas (não ligadas à equipe de pesquisadores):

Nome: _

Assinatura: _

Nome: _

Assinatura: _

APÊNDICE 5

Questionário para os participantes

1. Nome do entrevistador: _
2. Data: / / 2020. Hora de início: _ _
3. Nome do entrevistado: _
4. Profissão do entrevistado: _
5. Data de Nascimento: / /_
6. Idade: anos
7. Sexo: ( ) Feminino ( ) Masculino
8. Grau de escolaridade

( ) Ensino fundamental incompleto ( ) Ensino fundamental completo ( ) Ensino médio incompleto ( ) Ensino médio completo ( ) Ensino superior incompleto ( ) Ensino superior completo

9. Você teve algum sintoma de COVID19?

( ) Não **(pula para questão 15)** ( ) Sim

Se sim,

10. Data dos primeiros sintomas: / /2020
11. Quais sintomas?

( ) febre ( ) tosse seca ( ) cansaço
( ) dores e desconfortos ( ) dor de garganta
( ) diarreia ( ) conjuntivite ( ) dor de cabeça
( ) perda de paladar ou olfato ( ) erupção cutânea na pele ou descoloração dos dedos das mãos ou dos pés ( ) dificuldade de respirar ou falta de ar
( ) dor ou pressão no peito ( ) perda de fala ou movimento

( ) Outros: especificar _

12. Você realizou algum teste laboratorial para COVID19?

( ) Não **(pula para questão 15)** ( ) Sim

13. Quais testes laboratoriais?

( ) PCR : ( ) resultado Positivo ( ) resultado Negativo Data: / /2020

( ) Teste rápido IgM/IgG: ( ) resultado Positivo ( ) resultado Negativo Data: / /2020

( ) Teste rápido IgA/IgG: ( ) resultado Positivo ( ) resultado Negativo Data: / /2020

( ) Tomografia: % comprometimento do pulmão

14. Você foi hospitalizado com suspeita/confirmação de COVID-19?

( ) Não ( ) Sim: Por quantos dias?_ __

15. Qual (is) setor (es) do centro de saúde trabalha:

_____

16. Além deste local em que estamos fazendo a coleta, vc trabalha em algum outro que atenda ou atenderam pacientes com COVD19?

( ) Não, somente neste **(pula para questão 18)** ( ) Unidade básica de Saúde

( ) UPA ( ) Hospital **público** ( ) Hospital **privado** ( ) SAMU

( ) Consultório/Assistência domiciliar a pacientes ( ) Outro. Qual? _____

17. Trabalha em quantos lugares no total?

( ) 1 ( ) 2 ( ) 3 ( ) 4 ( ) 5 ou mais

18. Há quanto tempo você está prestando ou prestou assistência a pacientes com COVID-19? (em dias) _

19. Você apresenta alguma comorbidade como:

( ) Diabetes ( ) HAS (Hipertensão arterial sistêmica) ( ) Sobrepeso/Obesidade

( ) Cardiopatias ( ) Doença Renal ( ) Asma

20. Alguém da sua família **consanguínea** adoeceu (que mora ou não com você)?

( ) Não ( ) Sim: Quem?_

21. Alguém que mora com você adoeceu (sem **consanguinidade**)?

( ) Não ( ) Sim: Quantas pessoas? _____

ANEXOS

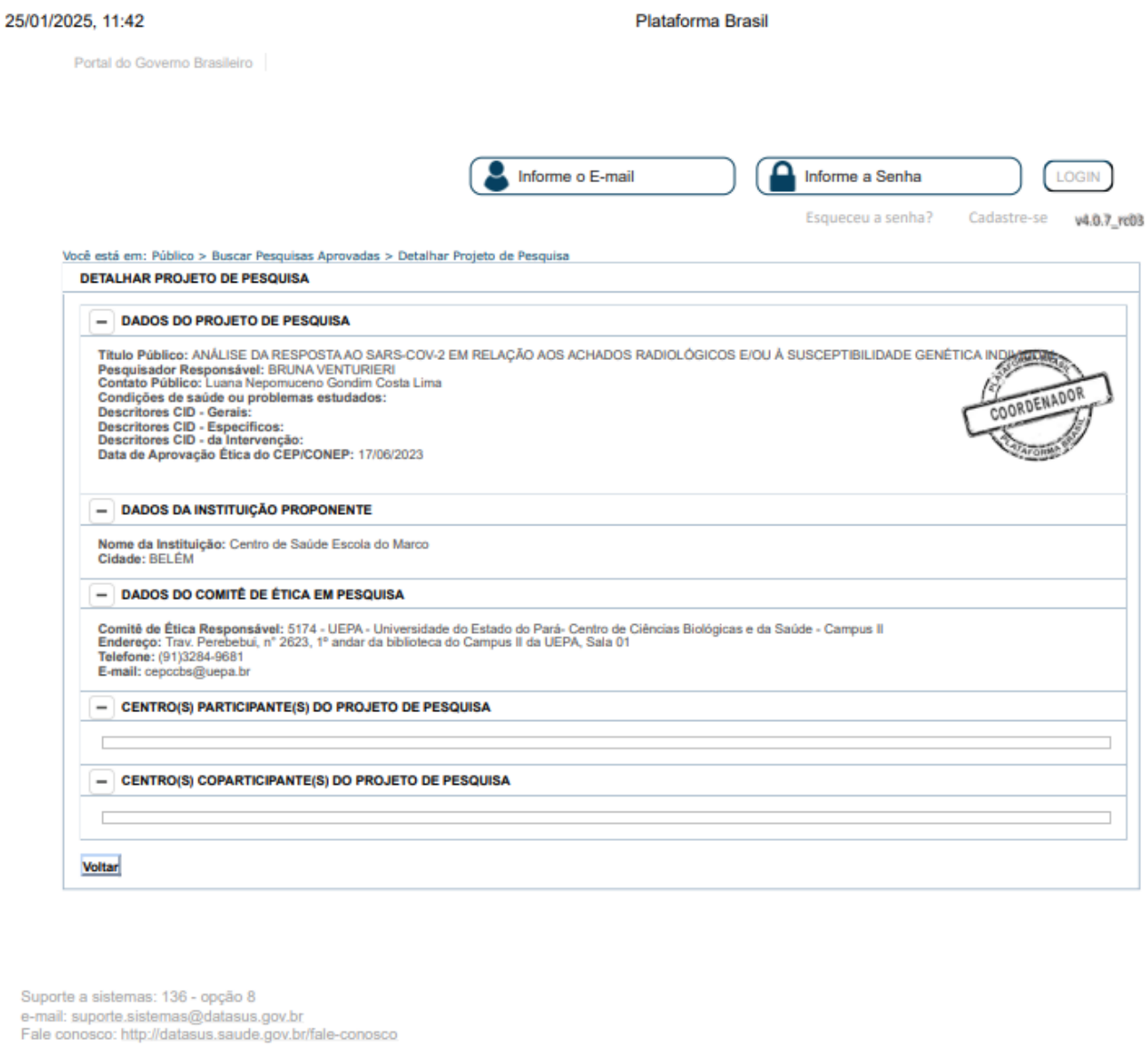
ANEXO 1 - Protocolo de extração QIAamp Blood Mini Kit

1. Pipetar 20 µL de Qiagen Protease em um tubo de microcentrífuga de 1,5 mL.
2. Homogeneizar os tubos com o sangue e transferir 200 µL da amostra de sangue (ponteira com
filtro) para o microtubo contendo a Qiagen Protease.
3. Adicione 200 µL de Tampão Al. Misture completamente por vórtex.
4. Incubar a 56°C por 25 min. em thermo shaker. Centrifugue brevemente o tubo de microcentrifugação de 1,5 mL para remover as gotas da tampa.
5. Adicionar 200 µL de etanol (96-100%). Misture bem por vórtex. Centrifugue brevemente o tubo
para remover as gotas da tampa.
6. Pipete a mistura na coluna de centrifugação QIAamp Mini (em um tubo de coleta de 2mL - utilizar
ponteira com filtro) e centrifugar a 8000 rpm durante 1 min. Descarte o tubo de coleta.
7. Colocar a coluna em novo tubo coletor e adicionar 500 µL de tampão AW1. Centrifugar por 1
min. a 8.000 rpm. Descartar o filtrado e o tubo coletor novamente. Ficar apenas com a coluna.
8. Colocar a coluna em novo tubo coletor e adicionar 500 µL de tampão AW2. Centrifugar por 3
min. a 13.000 rpm para deixar a membrana secar. Descartar o filtrado e o tubo coletor novamente. Ficar apenas com a coluna.
9. Coloque a coluna de centrifugação QIAamp Mini em um novo tubo de microcentrífuga de 1,5 mL
(não fornecido), adicione 90 µL de tampão AE e incube à temperatura ambiente (15-25°C) por 5
min. Centrifugar a 8000 rpm por 1 min para eluir o DNA.
10. O tubo coletor contém o DNA purificado. Tampar o microtubo e quantificar o DNA em
espectrofotômetro Nanodrop®.

11. Identificar o microtubo (tampa e corpo) indicando o CÓDIGO DA AMOSTRA, DATA e [DNA].

Inserir os dados no livro de protocolo e na planilha do Google Drive.

ANEXO 2 – Aceite no comitê de ética e pesquisa



25/01/2025, 11:42 Plataforma Brasil

Portal do Governo Brasileiro

Informe o E-mail | Informe a Senha | LOGIN

Esqueceu a senha? Cadastre-se v4.0.7_rc03

Você está em: Público > Buscar Pesquisas Aprovadas > Detalhar Projeto de Pesquisa

DETALHAR PROJETO DE PESQUISA

DADOS DO PROJETO DE PESQUISA

Título Público: ANÁLISE DA RESPOSTA AO SARS-COV-2 EM RELAÇÃO AOS ACHADOS RADIOLÓGICOS E/OU À SUSCEPTIBILIDADE GENÉTICA INDIVIDUAL
Pesquisador Responsável: BRUNA VENTURIERI
Contato Público: Luana Nepomuceno Gondim Costa Lima
Condições de saúde ou problemas estudados:
Descritores CID - Gerais:
Descritores CID - Específicos:
Descritores CID - da Intervenção:
Data de Aprovação Ética do CEP/CONEP: 17/06/2023

COORDENADOR

DADOS DA INSTITUIÇÃO PROPONENTE

Nome da Instituição: Centro de Saúde Escola do Marco
Cidade: BELÉM

DADOS DO COMITÊ DE ÉTICA EM PESQUISA

Comitê de Ética Responsável: 5174 - UEPA - Universidade do Estado do Pará- Centro de Ciências Biológicas e da Saúde - Campus II
Endereço: Trav. Perebebui, nº 2623, 1º andar da biblioteca do Campus II da UEPA, Sala 01
Telefone: (91)3284-9681
E-mail: cepccbs@uepa.br

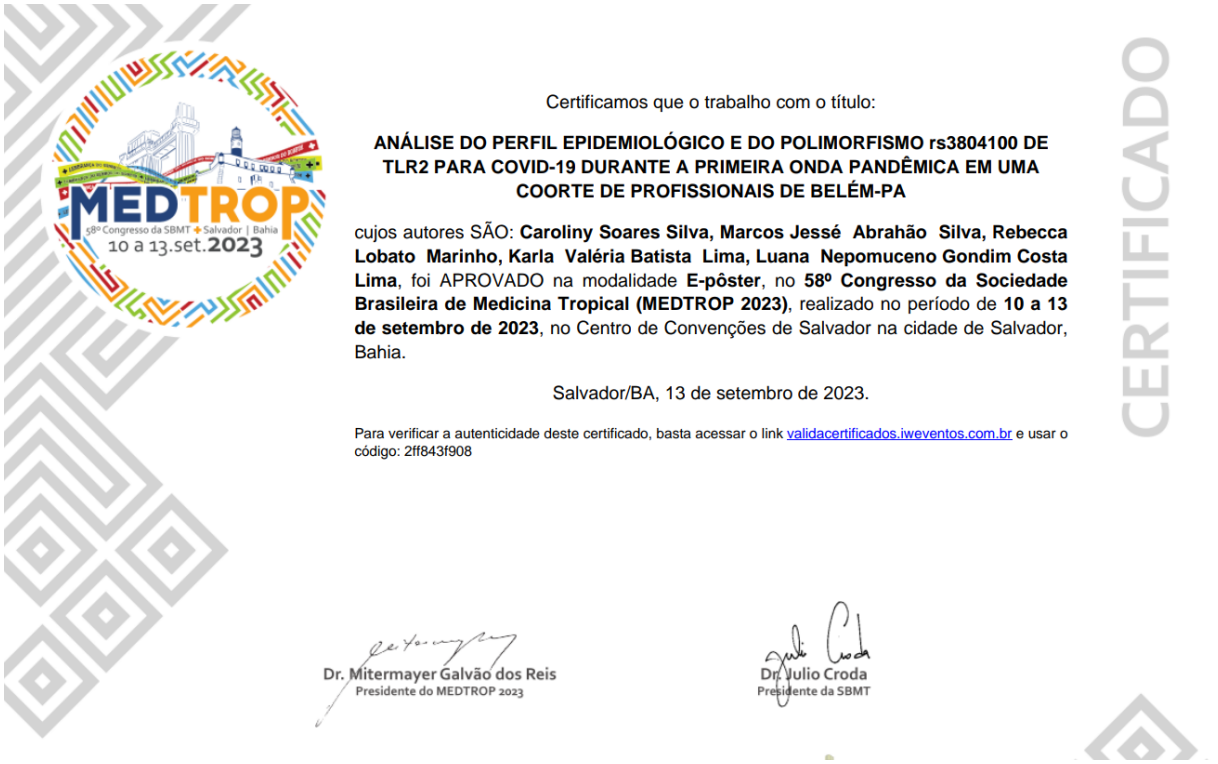
CENTRO(S) PARTICIPANTE(S) DO PROJETO DE PESQUISA

CENTRO(S) COPARTICIPANTE(S) DO PROJETO DE PESQUISA

Voltar

Suporte a sistemas: 136 - opção 8
e-mail: suporte.sistemas@datasus.gov.br
Fale conosco: http://datasus.saude.gov.br/fale-conosco

ANEXO 3 – Certificado de apresentação em dos resultados de polimorfismo na coorte no Congresso Brasileiro de Medicina Tropical em 2023



MEDTROP
58º Congresso da SBMT + Salvador | Bahia
10 a 13.set.2023

Certificamos que o trabalho com o título:

ANÁLISE DO PERFIL EPIDEMIOLÓGICO E DO POLIMORFISMO rs3804100 DE TLR2 PARA COVID-19 DURANTE A PRIMEIRA ONDA PANDÊMICA EM UMA COORTE DE PROFISSIONAIS DE BELÉM-PA

cujos autores SÃO: **Caroliny Soares Silva, Marcos Jessé Abrahão Silva, Rebecca Lobato Marinho, Karla Valéria Batista Lima, Luana Nepomuceno Gondim Costa Lima**, foi APROVADO na modalidade **E-pôster**, no **58º Congresso da Sociedade Brasileira de Medicina Tropical (MEDTROP 2023)**, realizado no período de **10 a 13 de setembro de 2023**, no Centro de Convenções de Salvador na cidade de Salvador, Bahia.

Salvador/BA, 13 de setembro de 2023.

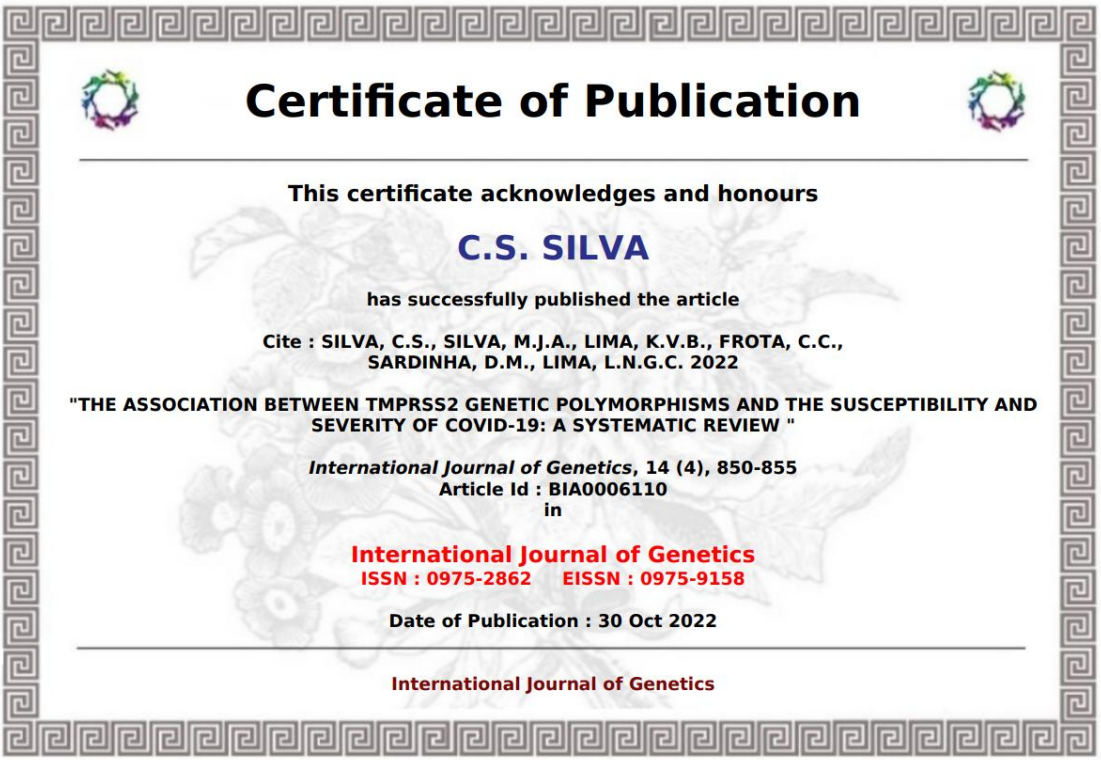
Para verificar a autenticidade deste certificado, basta acessar o link validacertificados.iweventos.com.br e usar o código: 2ff843f908

Dr. Mitermayer Galvão dos Reis
Presidente do MEDTROP 2023

Dr. Julio Croda
Presidente da SBMT

CERTIFICADO

ANEXO 4 – Certificado de publicação



Certificate of Publication

This certificate acknowledges and honours

C.S. SILVA

has successfully published the article

Cite : SILVA, C.S., SILVA, M.J.A., LIMA, K.V.B., FROTA, C.C., SARDINHA, D.M., LIMA, L.N.G.C. 2022

"THE ASSOCIATION BETWEEN TMPRSS2 GENETIC POLYMORPHISMS AND THE SUSCEPTIBILITY AND SEVERITY OF COVID-19: A SYSTEMATIC REVIEW "

International Journal of Genetics, 14 (4), 850-855
Article Id : BIA0006110
in

International Journal of Genetics
ISSN : 0975-2862 EISSN : 0975-9158

Date of Publication : 30 Oct 2022

International Journal of Genetics

ANEXO 5 – Certificado de publicação do Livro Comprehensive Bioinformatic Investigation of Exonic SNP’s of Candidate Genes of COVID-19

The work highlights the importance of functional characterization of single nucleotide polymorphisms (SNPs) associated with the risk for COVID-19. Using an in silico approach, exonic SNPs with potential functional impact and therapeutic value were identified, contributing to the development of personalized strategies to combat SARS-CoV-2 infection. This study, which uses in silico approaches to characterize exonic SNPs, highlights the importance of genomics in personalized medicine and the development of new therapeutic strategies.

SNPs on IFIH1, NOS3, IL-6R, TYK2, EFNA4, PLSCR1, IFNA10, SLC22A31, P2RX7, TLR7, IFNL1 and DDX58 genes were the main SNPs analyzed here, based on a literature search in scientific databases, regarding pathogenicity, structural and functional impacts on proteins and signaling pathways related to the high risk of COVID-19 disease. The investigation of SNPs associated with COVID-19 provides valuable insights into how genetic variations may influence susceptibility and severity of infection.

Functional analysis of synonymous and non-synonymous SNPs using advanced bioinformatics tools demonstrated that certain variants, such as sSNPs rs12252 in IFITM3 and rs731236 in VDR, have significant potential to impact mRNA structure and protein function, in addition to the nsSNP rs34536443 of TYK2 predicted to destabilize destabilize abilize the structure but also be capable of deleterious effects on the generated protein. The identification of deleterious SNPs suggests that some variants may predispose individuals to more severe forms of COVID-19, highlighting the need for genetic monitoring in vulnerable populations.

Marcos Jessé Abrahão Silva
Eliete Costa da Cruz
Rebecca Lobato Marinho

MJAS, ECC, RLM, DCL, LRTF, TNS, NCOA, KASN, CSS, LVMS, are Ph.D students on State University of Pará (UEPA), Brazil.
SKSB is graduate student at University of the Amazon (UNAMA), Brazil.
TPB and CCF are from Federal University of Ceará (UFC), Brazil.
LCSSS is from National Primate Center (CENP), Brazil.
YCR, TAFA, ECS, AJPG, LNGCL are from Evandro Chagas Institute, Brazil.

Comprehensive Bioinformatic Investigation of Exonic SNPs of Candidate Genes for COVID-19

9 789999 327732

ELIVA Press

ELIVA Press