UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FARMACOLOGIA E BIOQUÍMICA

RITA DE CÁSSIA CALDERARO COELHO

PERFIL MOLECULAR DAS VARIANTES POTENCIALMENTE ASSOCIADAS COM AS FORMAS SEVERAS DA COVID-19 EM POPULAÇÕES INDÍGENAS DA AMAZÔNIA

BELÉM – PA

2023

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FARMACOLOGIA E BIOQUÍMICA

PERFIL MOLECULAR DAS VARIANTES POTENCIALMENTE ASSOCIADAS COM AS FORMAS SEVERAS DA COVID-19 EM POPULAÇÕES INDÍGENAS DA AMAZÔNIA

Autora: RITA DE CÁSSIA CALDERARO COELHO

Orientador: Prof. Dr. NEY PEREIRA CARNEIRO DOS SANTOS

Documento de Dissertação apresentado ao Programa de Pós-graduação em Farmacologia e Bioquímica do Instituto de Ciências Biológicas da Universidade Federal do Pará como requisito para a obtenção do título de Mestre em Farmacologia e Bioquímica.

BELÉM – PA

2023

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBDSistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a)autor(a)

C672p Coelho, Rita de Cássia Calderaro.
Perfil Molecular das Variantes Potencialmente Associadas com as Formas Severas da COVID-19 emPopulações Indígenas da Amazônia / Rita de Cássia Calderaro Coelho. — 2023.
59 f. : il. color.

Orientador(a): Prof. Dr. Ney Pereira Carneiro dos SantosDissertação (Mestrado) - Universidade Federal do Pará,
Instituto de Ciências Biológicas, Programa de Pós- graduação em Farmacologia e Bioquímica, Belém, 2023.

1. COVID-19 . 2. SARS-CoV-2. 3. Infecção . 4.População indígena . 5. Perfil genético . I. Título.

CDD 576.5

DEDICATÓRIA/AGRADECIMENTOS

Antes de tudo, gostaria de agradecer à Deus que sempre foi o meu porto seguro me fazendo acreditar que eu seria capaz de concluir mais um passo na minha vida profissional. Agradeço aos meus pais, Herbet Luís Bastos Coelho e Rosália do Socorro Oliveira Calderaro Coelho, por todo o suporte na minha formação e por se fazerem tão presentes durante toda essa jornada.

Agradeço a minha avó Zuila Calderaro. Que mesmo sem imaginar me mostra dia após dia que a vida é cheia de propósitos e que um dia cada um deles será compreendido por nós.

Ao meu orientador, Prof. Dr. Ney Pereira Carneiro dos Santos e a Profª. Drª. Marianne Rodrigues Fernandes, por todos os aprendizados e direcionamentos dentro da pesquisa com conselhos que guardarei para a vida, serei sempre grata. Aos meus amigos do Laboratório de Medicina de Precisão, em especial aos que contribuíram de maneira mais direta com esse estudo: Lucas Favacho, Juliana Rodrigues, Kaio Aguiar, Maria Clara Calderaro, Laura Patrícia Gellen e Francisco Cezar Moraes, agradeço por toda a ajuda, conversas e aprendizados.

Agradeço também à minha família e aos meus amigos por sempre me apoiarem e ajudarem com palavras amigas em momentos de dificuldade, sendo fundamentais para mim.

À Universidade Federal do Pará (UFPA), ao Programa de Pós-graduação em Farmacologia e Bioquímica (FARMABIO) e ao Núcleo de Pesquisas em Oncologia (NPO) por terem plantado em mim o amor pela pesquisa científica. Agradeço também por todos os professores que contribuíram para a conclusão dessa etapa tão importante na minha vida acadêmica.

RESUMO

A doença do coronavírus 2019 (COVID-19) é uma doença causada pelo vírus SARS-CoV-2, podendo ser transmitida principalmente, por vias aéreas. Em razão da diversidade nas manifestações clínicas, a COVID-19 é considerada uma doença sistêmica, por comprometer vários órgãos e sistemas do corpo humano com sintomas que podem variar em assintomáticos a graves, podendo evoluir para óbito. Em virtude do aumento da cobertura vacinal e da disponibilização de novas vacinas no sistema de saúde, foi declarada o fim da emergência global da COVID-19. Apesar da atual estabilidade nas taxas de incidência e mortalidade pela doença, a sua progressão trouxe impactos negativos em todo o mundo. Muitos estudos de investigação do genoma amplo vem sendo desenvolvidos, como forma de compreender a relação de fatores genéticos com a gravidade da COVID-19. Contudo, a maior parte dessas investigações centram-se em países de origem europeia e asiática, que apresentam um perfil genético bem diferente das populações indígenas da Amazônia brasileira, evidenciando a necessidade de estudos que busquem maior compreensão do comportamento da COVID-19 nesses povos originários. Portanto, esse estudo tem como objetivo investigar variantes genéticas exclusivas da população indígena e compreender o perfil genético único dessa população, abrindo caminhos para estudos futuros em indígenas e contribuindo para a saúde desses povos.

Palavras-chave: COVID-19; SARS-CoV-2; Infecção; População indígena; Variantes e Perfil genético.

ABSTRACT

.Coronavirus disease 2019 (COVID-19) is a disease caused by the SARS-CoV-2 virus, which can be transmitted mainly by the air. Due to the diversity in clinical manifestations, COVID-19 is considered a systemic disease, because it compromises various organs and systems of the human body with symptoms that can vary from asymptomatic to severe, and can progress to death. Due to the increase in vaccination coverage and the availability of new vaccines in the health system, the end of the global COVID-19 emergency has been declared. Despite the current stability in incidence and mortality rates from the disease, its progression has brought negative impacts worldwide. Many studies of research of the broad genome have been developed as a way to understand the relationship of genetic factors with the severity of COVID-19. However, most of these investigations focus on countries of European and Asian origin, which have a very different genetic profile from the indigenous populations of the Brazilian Amazon, evidencing the need for studies that seek a greater understanding of the behavior of COVID-19 in these native peoples. Therefore, this study aims to investigate unique genetic variants of the indigenous population and understand the unique genetic profile of this population, opening paths for future studies in indigenous people and contributing to the health of these peoples.

.**Keywords:** COVID-19; SARS-CoV-2; Infection; Indigenous Population; Variants and Genetic Profile.

LISTA DE ILUSTRAÇÕES

LISTA DE SIGLAS E SÍMBOLOS

ICTV - Comitê Internacional de Taxonomia de Vírus

CoVs - Coronavírus

SARS-CoV-2 - Síndrome Respiratória Aguda Grave Coronavírus 2

RDB - *Domínio de Ligação ao Receptor (RBD)*

ECA-2 - Enzima Conversora de Angiotensina 2

TMPRSS2 - Serinoprotease Transmembrana 2

RE - Retículo Endoplasmático

OMS - *Organização Mundial da Saúde*

COVID-19 - Doença do Coronavírus 2019

ORFs - Quadros de Leitura Aberta (ORFs)

SDRA - Síndrome do Desconforto Respiratório Agudo

cTnI - Tromponina Cardíaca I (cTnI)

GWAS - *Genome-wide Association Studies*

SUMÁRIO

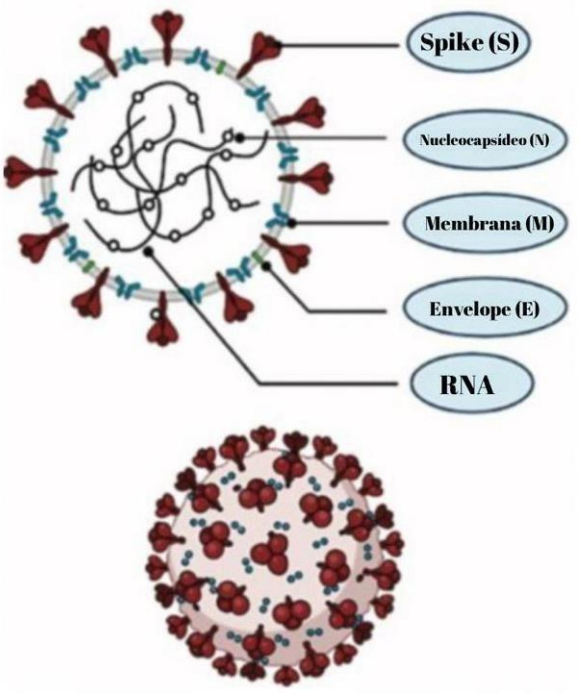
1. VISÃO INTEGRADORA DO PROBLEMA

1.1 CORONAVÍRUS - TAXONOMIA E GENOMA VIRAL

De acordo com a classificação do Comitê Internacional de Taxonomia de Vírus (ICTV), os coronavírus (CoVs) são vírus envelopados de RNA de fita simples positiva, pertencentes à família *Coronaviridae* e gêneros: *Alphacoronavirus*, *Betacoronavirus* *Deltacoronavirus* e *Gammacoronavirus* (Kim *et al.* 2020; ICTV, 2019; ICTV, 2020). Os casos iniciais da infecção por CoVs começaram em animais. Contudo, devido as altas e recorrentes taxas de mutações, esses vírus adquiriram a capacidade de infectar animais e seres humanos, através da chamada transmissão zoonótica (Zhou *et al.* 2020; Chan *et al.* 2013a; Chan *et al.* 2015b; Torres *et al.* 2022).

De maneira geral, os CoVs possuem uma morfologia semelhante entre si, sendo formados por proteínas do envelope (E), membrana (M) viral e proteína spike (S) (Figura 1). A entrada do vírus causador da Síndrome Respiratória Aguda Grave Coronavírus 2 (SARS-CoV-2) na célula hospedeira ocorre basicamente, pela ligação da membrana da célula infectada com a proteína S presente no vírus (Yao *et al.* 2020; Tortorici; Veesler, 2019; Hoffmann *et al.* 2020).

Figura 1 - Representação Morfológica do SARS-CoV-2.

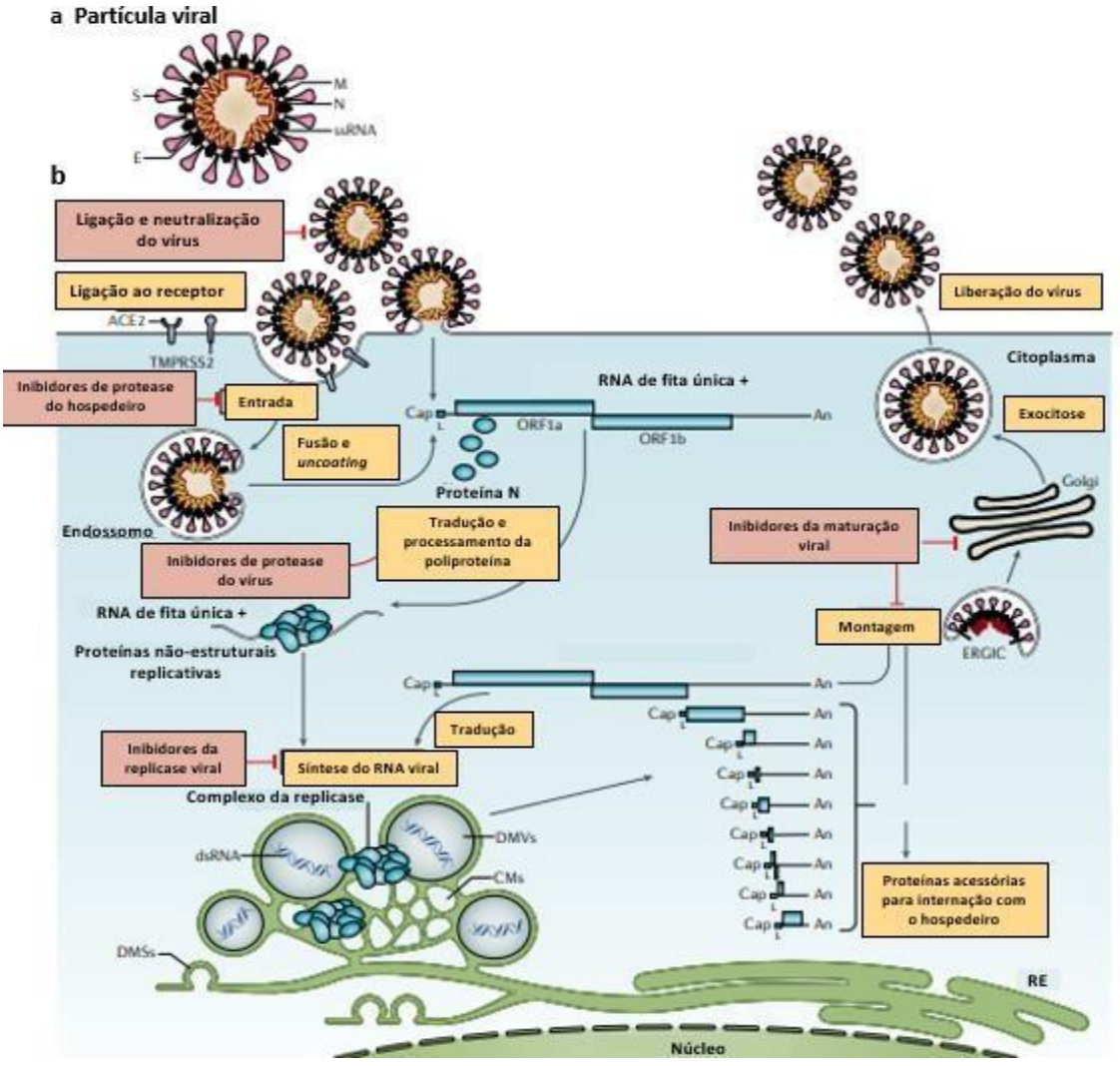


Fonte: Adaptado Torres *et al.* (2022).

Estudos sugerem que a interação entre célula/hospedeiro se inicia com o reconhecimento e ligação do Domínio de Ligação ao Receptor (RBD), encontrado no vírus e a Enzima Conversora de Angiotensina 2 (ECA-2), identificada na célula hospedeira (Hoffmann *et al.* 2020).

Posteriormente ao processo de ligação de RBD e ECA 2, ocorre uma clivagem dependente de proteases celulares (especialmente a Serinoprotease Transmembrana 2 – TMPRSS2). Tal processo de clivagem favorece a endocitose e a entrada do vírus na célula hospedeira (Hoffmann *et al.* 2020; Wang *et al.* 2020; Yan *et al.* 2020;Torres *et al.* 2022). Uma vez dentro da célula, a partícula viral libera seu material genético no citoplasma, o RNA liberado é então traduzido em duas poliproteínas (pp1a e pp1ab) que são clivadas por proteases virais (NSP3-Plpro e NSP5-Mpro) importantes para a síntese de proteínas não estruturais (Zhou *et al.* 2020;Fehr; Perlman, 2015) (Figura 2).

Figura 2 – Ciclo de replicação do SARS-CoV-2.



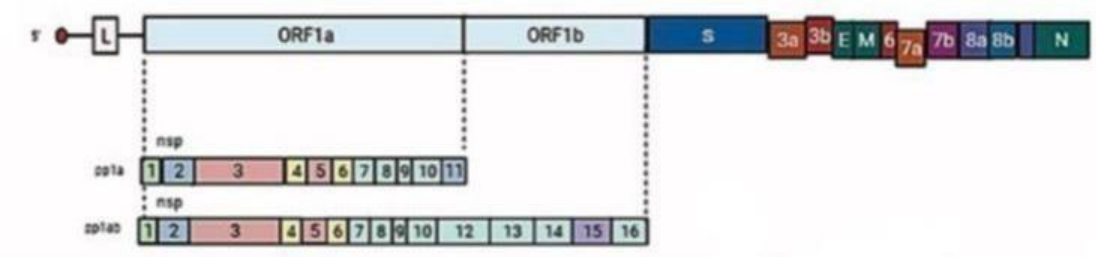
Fonte: Adaptado de V'kovski *et al.* (2021).

Ao final do ciclo de replicação, proteínas estruturais são inseridas no Retículo Endoplasmático (RE) através do compartilhamento ERGIC (contato entre RE e complexo de Golgi) (Jaimes, 2020; Torres *et al.* 2020) e novas partículas virais são produzidas e liberadas para fora da célula através do processo de exocitose.

À medida que os casos de infecção por SARS-CoV-2 começaram a se espalhar pelo mundo, a realidade da saúde pública global passou a ser cada vez mais crítica. Dada tal situação emergencial, a Organização Mundial da Saúde (OMS), declarou a pandemia da doença do coronavírus 2019 (COVID-19) (Chan, 2003c; Dhama *et al.* 2020; Cucinotta; Vanelli *et al.* 2020). A COVID-19 é uma doença causada pelo SARS-CoV-2, transmitida através de vias aéreas e com uma manifestação clínica bastante extensa, sendo considerada uma doença sistêmica por comprometer o funcionamento de vários órgãos do corpo humano (Driggin *et al.* 2020; Terpos *et al.* 2020).

O genoma do vírus causador da COVID-19 é formado por: 13 quadros de leitura aberta (ORFs); 4 proteínas estruturais: spike (S), nucleocapsídeo (N), membrana (M) e envelope (E); proteínas não estruturais (NSP1, NSP2, NSP3, NSP4, NSP5, NSP6, NSP7, NSP8, NSP9, NSP10, NSP12, NSP13, NSP14, NSP15 e NSP16); e proteínas acessórias (3a, 3b, 6, 7a, 7b, 8b, 9b e ORF14) (Figura 3) (Lu *et al.* 2020; Wu *et al.* 2020; Torres *et al.* 2022).

Figura 3 – Genoma do SARS-CoV-2.



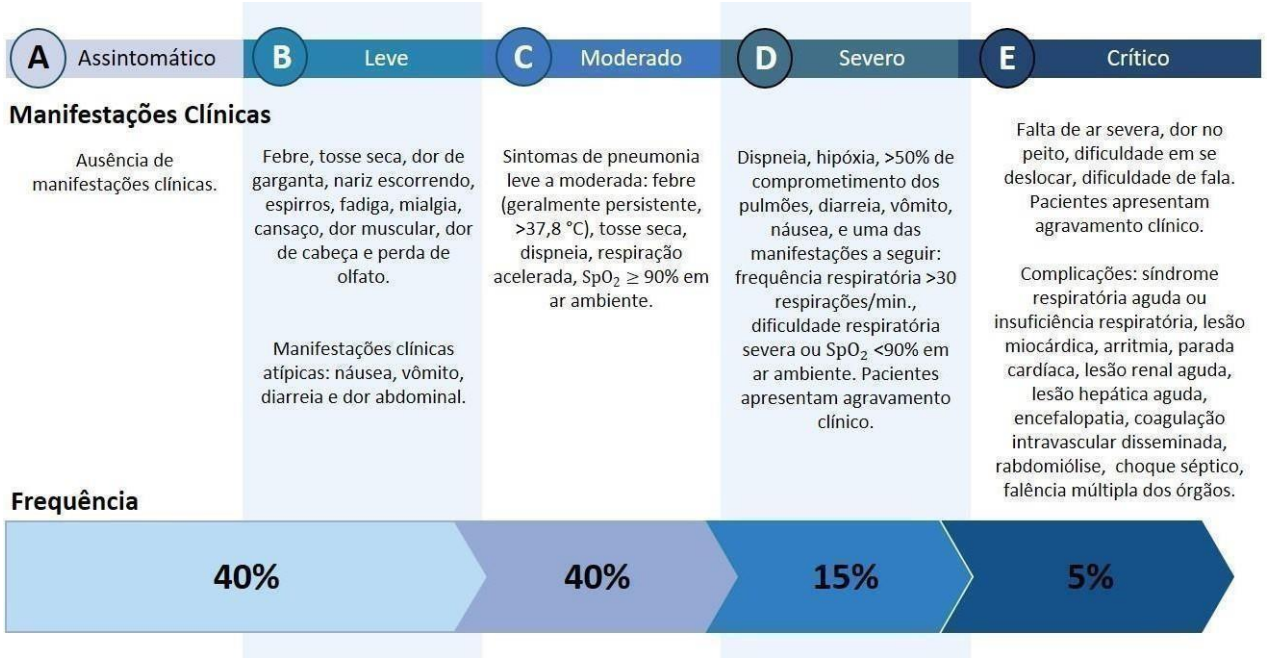
Fonte: Adaptado de Torres *et al.* (2022).

1.2 VARIAÇÕES CLÍNICAS DA COVID-19

Os sintomas da COVID-19 variam em assintomáticos, leves, moderados, severos ou críticos, podendo evoluir para óbito (Figura 4) (Fricke-Galindo; Falcan-Valencia, 2021; Huang *et al.* 2020). Mais de 40% dos pacientes com a COVID-19 apresentam a Síndrome do Desconforto Respiratório Agudo (SDRA), especialmente aqueles que têm algum tipo de comorbidade (Wu *et al.* 2020). Quadros de linfopenia também são bastante associados com os casos mais graves da doença, evidenciando

o quanto a redução de leucócitos está relacionada com os casos letais da COVID-19 (Henry, 2020).

Figura 4 – Manifestações clínicas da COVID-19.



Fonte: Adaptado de Fricke-Galindo; Falcan-Valencia, (2021).

A presença de biomarcadores cardíacos também indica casos graves da doença. Em um estudo foi demonstrado que a concentração de Tromponina Cardíaca I (cTnI) aumentou em pacientes diagnoticados com a COVID-19, quando comparados com aqueles que desenvolveram sintomas mais leves (Lippi; Lavie; Sanchis-Gomar, 2020). Em razão ao amplo espectro clínico da doença, pacientes com idade avançada e/ou que apresentam doenças como diabetes, hipertensão, comprometimentos cardiovasculares, dentre outras comorbidades, podem desenvolver um quadro mais grave da COVID-19 (Huang *et al.* 2020).

No entanto, ainda que tais fatores sejam associados com a severidade e mortalidade da doença, essas condições não podem ser a única explicação para a evolução clínica da COVID-19. Hábitos de vida, influência da alimentação, fatores culturais e genéticos também são fatores bastante influentes na resposta individual a doença (Park *et al.* 2020; Fricke-Galindo; Falcan-Valencia, 2021). No que concerne isso, é de extrema importância o monitoramento de biomarcadores dos pacientes antes e após a internação para auxiliar na compreensão da progressão da COVID-19 (Lippi; Lavie; Sanchis-Gomar, 2020).

1.3 EPIDEMIOLOGIA

1.3.1 Mundial

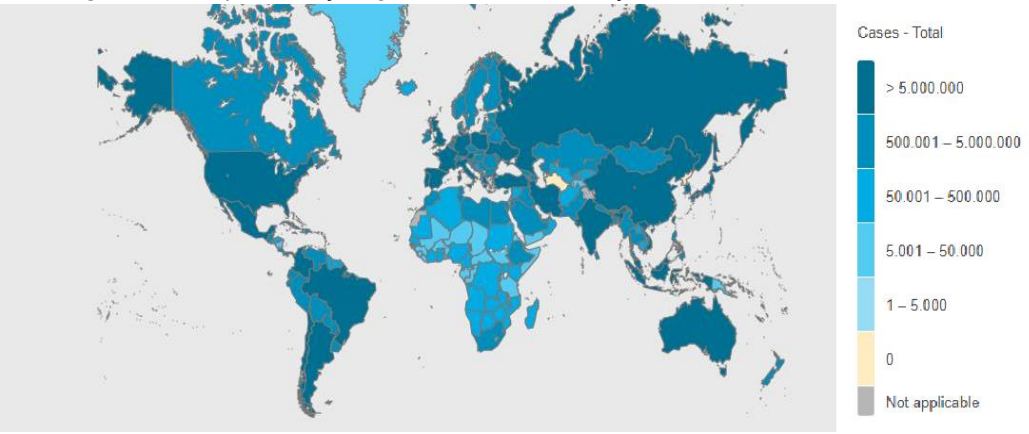
Segundo a OMS (2004), casos de infecção por SARS foram identificados em 2002, na província chinesa de Guangdong, levando a óbito 916 pessoas em vários países. Anos mais tarde, casos de infecção por SARS-CoV em humanos foram confirmados com 774 mortes (Drosten *et al.* 2003; Lau *et al.* 2003).

Em dezembro de 2019, casos de pneumonia ocasionados por um agente etiológico até então desconhecido, começaram a ser reportados. Pouco tempo depois foram confirmados mais de 414.179 casos da COVID-19 reportados em 197 países, sendo declarada a pandemia causada pelo SARS-CoV-2, a COVID-19 (Baloch *et al.* 2020).

É incontestável os impactos negativos gerados em decorrência da pandemia. No entanto, em virtude do aumento da cobertura vacinal e da disponibilização de um maior número de vacinas no sistema de saúde, recentemente foi declarado pela OMS, o fim da emergência global da COVID-19 (Luo *et al.* 2023). A vacinação da população tornou possível a redução das taxas de internação e mortalidade pela doença, contribuindo para a atual estabilidade dos casos no mundo todo (Rabaan *et al.* 2020; Woo *et al.* 2005).

Atualmente, de acordo com dados da OMS, até o dia 27 de setembro de 2023, houveram 770.875.433 casos confirmados da COVID-19, com 6.959.316 óbitos. No entanto, são relatados semanalmente redução no número de mortalidade pela doença em níveis mundiais (Figura 5).

Figura 5 – Representação gráfica atual da situação da COVID-19 no mundo.



Fonte: Adaptado da WHO, 2023.

1.3.2 Brasil

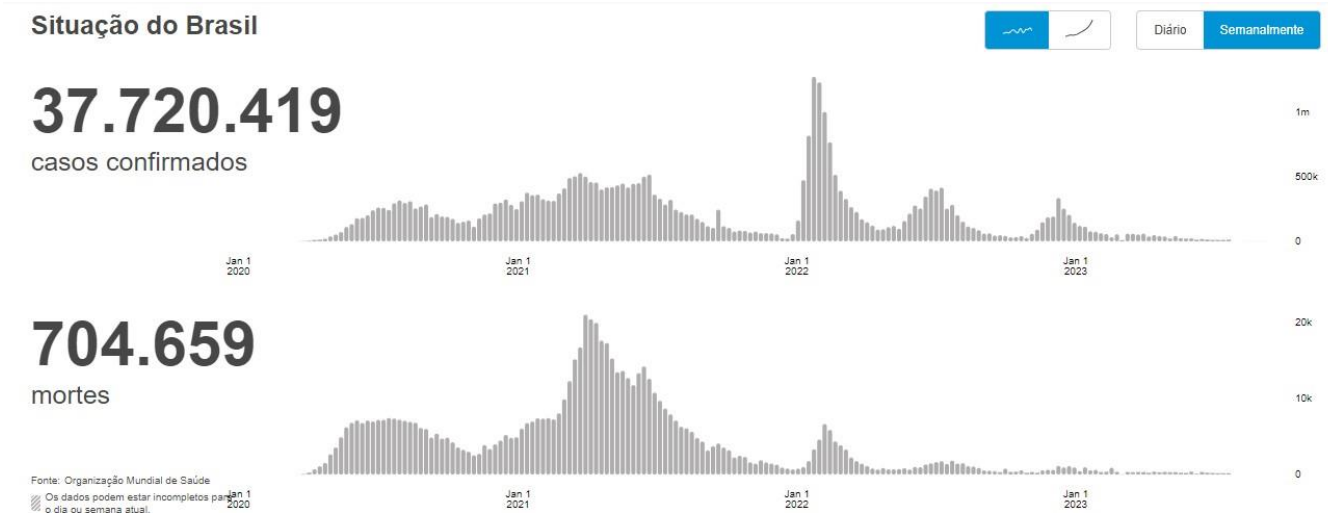
No dia 26 de fevereiro de 2020, em São Paulo, foi confirmado o primeiro caso da COVID-19 no Brasil. Um indivíduo do sexo masculino que tinha retornado de uma viagem à Itália tinha entrado em contato com o vírus (Araújo *et al.* 2020). Até a confirmação desse primeiro caso no Brasil, já haviam 81.109 pessoas infectadas em 38 países (Rodriguez-Morales *et al.* 2020). A partir do primeiro caso confirmado da doença no país, o número de pacientes diagnosticados com a COVID-19 passou a ser cada vez mais crescente, chegando em agosto de 2020 a 3.846.153 casos confirmados e 120.462 óbitos (WHO, 2023).

Devido a rápida progressão, fez-se necessário a adoção de medidas emergenciais como forma de controlar os casos novos da doença. Tais estratégias preventivas visavam isolar a pessoa infectada e também realizar diagnóstico precoce da COVID-19. Como forma de reduzir a transmissão viral também eram aconselhadas práticas de higiene como lavagem correta das mãos, evitar contato com mucosas, dentre outras. O fechamento de fronteiras, escolas e outros serviços em geral, também foram algumas outras medidas praticadas em vários países como forma de minimizar o impacto da pandemia (Güner, 2020).

Apesar da implementação dessas medidas terem demonstrado um efeito positivo no controle da pandemia retardando a transmissão da doença, isso não foi o suficiente para o seu controle. Tanto no Brasil quanto no mundo, o avanço nos casos da COVID-19 e aumento das subnotificações, foram cada vez mais frequentes. Contudo, atualmente esses dados apresentam-se estáveis em níveis globais.

De acordo com dados coletados de janeiro à setembro de 2023, o Brasil apresentou 37.720.419 casos confirmados por COVID-19 com 704.659 óbitos (Figura 6). Ainda que esses números sejam bastante expressivos, atualmente, no dia 25 de setembro de 2023, o Brasil não demonstrou aumento no número de casos e mortalidade por COVID-19.Tais fatos evidenciam que a situação atual do país em relação a COVID-19, é bastante positiva (WHO, 2023).

Figura 6 – Situação atual do Brasil em relação a COVID-19.



Fonte: Adaptado da WHO, 2023.

1.4 POPULAÇÕES INDÍGENAS

A sub-representação das populações indígenas em estudos genômicos sobre a COVID-19 representa uma escassez de informação em relação a resposta individual sobre o comportamento da doença dentro destes grupos populacionais. Consequentemente, devido ao perfil genético único das populações indígenas, pouco se sabe a respeito das influências de variantes genéticas presentes nesses grupos que podem interferir na resposta individual à doença em cada indivíduo, de acordo comsuas particularidades.

Presentes há séculos no território brasileiro, a população indígena apresenta uma forte contribuição tanto no processo histórico do Brasil quanto na identidade cultural, territorial e no que diz respeito a influência genômica individual. Dados do Censo Demográfico de 2010 demonstram um crescimento de pessoas autodeclaradas indígenas nas regiões Norte, Nordeste e Centro-Oeste. Adicionalmente, apesar da presença de populações indígenas em todo o Brasil, o Norte do país apresenta o maior percentual de autodeclarados indígenas, com aproximadamente 37,4% da totalidade de povos indígenas no país (IBGE, 2010).

Pouco se sabe a respeito da saúde indígena no Brasil, especificamente se tratando de estudos de associação de GWAS *(Genome-wide Association Studies)* relacionados com a investigação de variantes genéticas envolvidas com doenças infecciosas. Além disso, outros estudos demonstram a mudança nos hábitos alimentares da população indígena, por meio da introdução de alimentos industrializados e ultra processados, bem como o consumo de bebidas alcoólicas,

frequentemente relacionados com doenças como obesidade e diabetes (Alvim *et al.* 2014; Arnaiz- Villena *et al.* 2013; Barbosa *et al.* 2019). Tais doenças podem acarretar em problemas de saúde mais graves que venham a comprometer o perfil genético de cada indivíduo e, consequentemente, a maneira como a COVID-19 pode ser manifestada.

São escassos os dados na literatura que abordam a relação dos grupos indígenas com a COVID-19, principalmente estudos que demonstram a associação de variantes genéticas com a severidade da doença dentro dessas populações. Dentro desse contexto, recentemente uma investigação desenvolvida em nosso grupo de pesquisa investigou variantes genéticas nos genes *SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6, XCR1* e *ABO* potencialmente envolvidos com as formas graves da COVID-19 nas populações indígenas: Asurini do Tocantins, Asurini do Xingu, Araweté, Arara, Juruna, Awa-Guajá, Kayapó/ Xikrin, Munduruku, Karipuna, Phurere, Wajãpi e Zo'é (Pastana *et al.* 2022). O trabalho demonstrou que o gene *ABO* foi o que apresentou a maioria das variantes com diferenças significativas, o que sugere que a diferença no perfil genômico e a identificação de novas variantes na população indígena pode influenciar no desenvolvimento das formas graves da doença.

No que concerne isso, é imprescindível maiores pesquisas em populações indígenas como forma de compreender o perfil genético único desses grupos. Sendo assim, é fundamental o desenvolvimento de estudos genômicos que visem investigar variantes genéticas com potencial associação a gravidade da COVID-19, que sejam exclusivas desse grupo e que contribuam para o desenvolvimento da medicina personalizada respeitando a individualidade da população indígena.

1.4.1 Estudos genômicos em populações indígenas da Amazônia

Apesar do fim da emergência global da pandemia, a COVID-19 atingiu proporções mundiais desanimadoras. A progressão de casos da doença acarretou em grandes impactos negativos em todo o mundo. Tendo em vista tais impactos mundiais da pandemia, estudos de investigação de genoma amplo são fundamentais para a compreensão do desenvolvimento da COVID-19 e suas consequências.

Estudos de GWAS identificaram variantes genéticas com forte relação a COVID-19 grave (Ellinghaus *et al*. 2020; Shelton *et al.* 2021). Um estudo demonstrou forte associação dos rs11385942 (locus 3p21.31) e rs657152 (locus 9q34.2) com a forma

mais grave da doença. Além disso, o estudo também demonstrou a presença dos genes *SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6* e *XCR1* no locus 3p21.31 e do grupo sanguíneo *ABO* (locus 9q34.2) (Ellinghaus *et al.* 2020). Amostras de europeus, latinos e africanos foram analisadas em outro estudo e constatou-se que o grupo sanguíneo *ABO* aumentou a suscetibilidade a gravidade da COVID-19. Entretanto, diferente dos outros grupos sanguíneos analisados no estudo,o grupo O foi sugerido como um fator protetor contra a doença (Shelton *et al.* 2021).

Em outro estudo recente com pacientes de 34 hospitais da Espanha novos genes foram associados com a gravidade da COVID-19 (*AQP3, ARHGAP27, ELF5, IFNAR2, LIMD1, OAS1* e *UPK1A*). Neste estudo, também foi abordado o comportamento da doença nos sexos masculino e feminino (Cruz *et al.* 2022). Os mesmos genes presentes na pesquisa de Cruz *et al.* 2022 também foram selecionados para o presente estudo, mas com amostras da população indígena da Amazônia brasileira.

Dessa forma, estudos genômicos que busquem compreender a variabilidade genética das populações indígenas e que sejam capazes de investigar variantes genéticas exclusivas desse grupo refletem positivamente para estudos futuros que comprovem o impacto dessas variantes em pacientes com a COVID-19 em populações indígenas, contribuindo para a saúde pública desses povos.

2. ARTIGO 1: O Perfil Genômico Associado ao Risco de Formas Graves de COVID-19 em Nativos Americanos da Amazônia

Journal of *Personalized Medicine*

MDPI

Article

The Genomic Profile Associated with Risk of Severe Forms of COVID-19 in Amazonian Native American Populations

Lucas Favacho Pastana [1], Thays Amâncio Silva [1], Laura Patrícia Albarello Gellen [1], Giovana Miranda Vieira [1], Letícia Almeida de Assunção [1], Luciana Pereira Colares Leitão [1], Natasha Monte da Silva [1], Rita de Cássia Calderaro Coelho [1], Angélica Leite de Alcântara [1], Lui Wallacy Morikawa Souza Vinagre [1], Juliana Carla Gomes Rodrigues [1], Diana Feio da Veiga Borges Leal [1], Marianne Rodrigues Fernandes [1], Sandro José de Souza [2,3,4], José Eduardo Kroll [2], André Mauricio Ribeiro-dos-Santos [5], Rommel Mario Rodríguez Burbano [1], João Farias Guerreiro [5], Paulo Pimentel de Assumpção [1], Ândrea Campos Ribeiro-dos-Santos [5], Sidney Emanuel Batista dos Santos [1] and Ney Pereira Carneiro dos Santos [1,*]

[1] Laboratório do Núcleo de Pesquisa em Oncologia, Universidade Federal do Pará, Belém 66073-000, Brazil; lucas.pastana@ics.ufpa.br (L.F.P.); thaysamnc@gmail.com (T.A.S.); laura.patricia.agellen@hotmail.com (L.P.A.G.); giovanamirandav@gmail.com (G.M.V.); leticiaalmeidaenf96@gmail.com (L.A.d.A.); colares.luciana@gmail.com (L.P.C.L.); ntshmonte@gmail.com (N.M.d.S.); rccalderarocoelho@gmail.com (R.d.C.C.C.); angelica.alcantara99@gmail.com (A.L.d.A.); luivinagre@gmail.com (L.W.M.S.V.); julianacrodrigues@gmail.com (J.C.G.R.); dianafeio@hotmail.com (D.F.d.V.B.L.); fernandesmr@yahoo.com.br (M.R.F.); rommelburbano@gmail.com (R.M.R.B.); assumpcaopp@gmail.com (P.P.d.A.); sidneysantosufpa@gmail.com (S.E.B.d.S.)

[2] Instituto do Cérebro, Universidade Federal do Rio Grande do Norte, Natal 59076-550, Brazil; sandro@neuro.ufrn.br (S.J.d.S.); jkpenga@gmail.com (J.E.K.)

[3] BioME, Universidade Federal do Rio Grande do Norte, Natal 59078-400, Brazil

[4] Institute of Systems Genetics, West China Hospital, University of Sichuan, Chengdu 610041, China

[5] Laboratório de Genética Humana e Médica, Universidade Federal do Pará, Belém 66075-110, Brazil; andremrsantos@gmail.com (A.M.R.-d.-S.); joao.guerreiro53@gmail.com (J.F.G.); akelyufpa@gmail.com (Â.C.R.-d.-S.)

* Correspondence: npcsantos.ufpa@gmail.com; Tel.: +55-(91)-98107-0850

Abstract: Genetic factors associated with COVID-19 disease outcomes are poorly understood. This study aimed to associate genetic variants in the *SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6*, *XCR1*, and *ABO* genes with the risk of severe forms of COVID-19 in Amazonian Native Americans, and to compare the frequencies with continental populations. The study population was composed of 64 Amerindians from the Amazon region of northern Brazil. The difference in frequencies between the populations was analyzed using Fisher's exact test, and the results were significant when $p \leq 0.05$. We investigated 64 polymorphisms in 7 genes; we studied 47 genetic variants that were new or had impact predictions of high, moderate, or modifier. We identified 15 polymorphisms with moderate impact prediction in 4 genes (*ABO*, *CXCR6*, *FYCO1*, and *SLC6A20*). Among the variants analyzed, 18 showed significant differences in allele frequency in the NAM population when compared to others. We reported two new genetic variants with modifier impact in the Amazonian population that could be studied to validate the possible associations with COVID-19 outcomes. The genomic profile of Amazonian Native Americans may be associated with protection from severe forms of COVID-19. This work provides genomic data that may help forthcoming studies to improve COVID-19 outcomes.

Keywords: COVID-19; gene; risk factor; genetic variant

Citation: Pastana, L.F.; Silva, T.A.; Gellen, L.P.A.; Vieira, G.M.; de Assunção, L.A.; Leitão, L.P.C.; da Silva, N.M.; Coelho, R.d.C.C.; de Alcântara, A.L.; Vinagre, L.W.M.S.; et al. The Genomic Profile Associated with Risk of Severe Forms of COVID-19 in Amazonian Native American Populations. *J. Pers. Med.* **2022**, *12*, 554. https://doi.org/10.3390/jpm12040554

Academic Editors: Patrick Plaisance and Anthony Chauvin

Received: 25 January 2022
Accepted: 15 February 2022
Published: 1 April 2022

1. Introduction

The coronavirus disease 2019 (COVID-19) outbreak started when a few patients were hospitalized with acute respiratory distress syndrome in December 2019. At the end of

January 2020, a total of 1975 COVID-19 cases were confirmed in China, with a total of 56 deaths [1]. The new infection spread worldwide, and the World Health Organization (WHO) declared COVID-19 a pandemic on 11 March 2020 [2]. Globally, the number ofconfirmed cases of COVID-19 has reached almost 386,548,962, including 5,705,754 deaths asof 6 February 2022 [3]

All individuals are susceptible to COVID-19 infection; however, the severity of the disease varies significantly between individuals and populations. There are many host, viral, and environmental factors contributing to the COVID-19 phenotype [4]; however, the genetic factors associated with COVID-19 disease outcomes are poorly understood. The discovery of human genetic factors associated with this disease's severity would be in- valuable in identifying high-risk groups, and would enable the stratification of individuals according to risk in order to guide personalized prevention and therapeutics [5].

Infectious diseases continue to disproportionately affect indigenous peoples and admixture populations with Amerindian ancestry [6–10], and COVID-19 has reached indigenous communities [11–16]. This population has a particular genetic vulnerability to infection due to different frequencies of alleles in immune system genes [17]. Their high genetic homozygosity has been suggested to be a consequence of a serial founder effect, compounded by successive generations of inbreeding [18]. This genetic factor may result in a significant loss of diversity and have consequences on health and performance [8].

In genome-wide association studies (GWASs), the identification of potential genetic factors associated with the development of COVID-19 has been investigated. The first GWAS analysis with 1980 patients with COVID-19 identified two loci associated with the most severe forms of COVID-19: one locus was 3p21.31, which includes the genes *SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6*, and *XCR1*, while the other was 9q34.21, including the *ABO* blood group. These results might explain the heterogeneity of the disease [19].

The *CXCR6*, *CCR9*, and *XCR1* genes are chemokine receptors, and directly participate in the functioning of cells of the immune system and the expression of interleukins. Other selected genes have more heterogeneous actions. The *LZTFL1* gene plays a role in intracellular signaling actions. The *FYCO1* gene is involved in the transport of autophagic vesicles, the *SLC6A20* gene activates virus adhesion co-receptors in the cell, and the *ABO* gene is related to glycosylation of the H antigen for the formation of blood group variability [19]. Subsequently, Shelton et al. identified a strong association between blood type and COVID-19 diagnosis. Moreover, variants on chromosome 3p21.31 were strongly associated with COVID-19 outcome severity [20]. The present study investigated genetic variants in the *SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6*, *XCR1*, and *ABO* genes that are potentially related to severe forms of COVID-19 in Amazonian Native American populations, and compared their frequencies with continental populations.

2. Materials and Methods

2.1. Study and Reference Populations

All participants in the study and their ethnic group leaders signed written informed consent. The recruitment period was before the COVID-19 pandemic, from September 2017 to December 2018. The study was approved by the National Committee for Ethics in Research (CONEP) and the Research Ethics Committee of the UFPA Tropical Medicine Center under CAAE number 20654313.6.0000.5172, and by the Research Ethics Committee of the UFPA under project 123/98.

The Amazonian Native American (NAM) cohort was composed of 64 Amerindians from the Amazon region of northern Brazil. The NAM population was healthy and did not share family relationships. The genetic ancestry was obtained through a panel of 61 ancestry-informative markers (AIMs), which were used for estimating individual ancestry and admixture from three continents (European, African, and Amerindian) in three multiplex PCR reactions [7,21,22]. The amplicons were analyzed by electrophoresis using the ABI Prism 3130 sequencer and the GeneMapper ID v.3.2 software. The individual proportions were estimated using STRUCTURE v.2.3.3.

The allele frequencies of the NAM population were obtained directly by gene counting and compared with continental populations (available at http://www.1000genomes.org). The 1000 Genomes project dataset corresponds to full-length DNA sequences from 2504 hu- man individuals that include 26 population groups clustered into 5 larger population groups (African: AFR; American: AMR; East Asian: EAS; European: EUR; and South Asian: SAS). These populations include 661 individuals of African (AFR), 347 of American (AMR), 504 of East Asian (EAS), 503 of European (EUR), and 489 of South Asian (SAS) descent.

For the samples with European, East Asian, and South Asian ancestry, populations across the geographic range had ~1% FST; populations from Africa were related to the Yoruba and, therefore, not a comprehensive representation of Africa; for populations in the Americas, the samples were from two populations with primarily African and European ancestry— people with African Ancestry in the southwest USA (ASW), and those of Afro-Caribbean descent in Barbados (ACB)—and four populations (people with Mexican Ancestry in Los Angeles, CA, USA (MXL), Colombians in Medellin, Colombia (CLM), Puerto Ricans in Puerto Rico (PUR), and Peruvians in Lima, Peru (PEL)) with a wide range of European, African, and Indigenous American ancestry were chosen to represent the wide variation in ancestry proportions observed in North, Central, and South America.

2.2. *Extraction of DNA and Preparation of the Exome Library*

The DNA extraction was performed via the phenol–chloroform method [23]. The quantification and integrity of genetic material were analyzed using a NanoDrop 8000 spec- trophotometer (Thermo Fisher Scientific Inc., Wilmington, DE, USA) and electrophoresis in 2% agarose gel, respectively.

The exome libraries were prepared using the Nextera Rapid Capture Exome (Illumina®, San Diego, CA, USA) and SureSelect Human All Exon V6 (Agilent, Santa Clara, CA, USA) kits. The sequencing reactions were run on the NextSeq 500® platform (Illumina®, San Diego, CA, USA) using the NextSeq 500 High-Output v2 300 cycle kit (Illumina®, San Diego, CA, USA).

2.3. *Bioinformatic Analysis*

The quality of the FASTQ reads was analyzed (FastQC v.0.11—http://www.bioinformatics.babraham.ac.uk/projects/fastqc/; accessed on 20 January 2022), and the samples were filtered to eliminate low-quality readings (fastx_tools v.0.13—http://hannonlab.cshl.edu/fastx_toolkit/; accessed on 20 January 2022). The sequences were mapped and aligned with the reference genome (GRCh38) using the BWA v.0.7 tool (http://bio-bwa.sourceforge.net/; accessed on 20 January 2022). Following this alignment with the reference genome, the file was indexed and sorted (SAMtools v.1.2—http://sourceforge.net/projects/samtools/; accessed on 20 January 2022). Subsequently, the alignment was processed for duplicate PCR removal (Picard Tools v.1.129—http://broadinstitute.github.io/picard/; accessed on 20 January 2022), mapping quality recalibration, and local realignment (GATK v.3.2—https://www.broadinstitute.org/gatk/; accessed on 20 January 2022). The results were processed in order to determine the variants from the reference genome (GATK v.3.2). The analysis of the variant annotations was carried out using the ViVa1 (Viewer of Variants) software developed by the Federal University of Rio Grande do Norte (UFRN)'s bioinformatics team. The databases and their versions used for variant annotations were SnpEff v.4.3.T, Ensembl Variant Effect Predictor (Ensembl release 99), and ClinVar (v.2018- 10). For in silico prediction of pathogenicity, we used SIFT (v.6.2.1), PolyPhen-2 (v.2.2), LRT (November 2009),Mutation Assessor (v.3.0), Mutation Taster (v. 2.0), FATHMM (v.2.3), PROVEAN (v.1.1.3), MetaSVM (v1.0), M-CAP (v1.4), and FATHMM-MKL (http://fathmm.biocompute.org.uk/about.html; accessed on 20 January 2022). More information about bioinformatic analyses is described in the works of Rodrigues et al. (2020) and Ribeiro-dos-Santos et al. (2020) [22,24].

2.4. Statistical Analyses

The difference in frequencies between the populations was analyzed using Fisher's exact test, and the results were significant when $p \leq 0.05$. The interpopulation variability of the polymorphisms was assessed using Wright's fixation index (FST). These analyses were performed using RStudio v.4.1.0.

2.5. Selection of Variants

Seven genes (*SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6*, *XCR1*, and *ABO*) used in recent GWAS studies were selected. These studies identified genes at the loci 3p21.31 and 9q34.21 as being likely related to disease severity [19,20,25]. For subsequent analyses, the selection of variants was based on three main criteria: (a) a minimum of 10 reads of coverage (fastx_tools v.0.13-http://hannonlab.cshl.edu/fastx_toolkit/; accessed on 20 January 2022), (b) the variant should have an allele frequency described in all continental populations from the 1000 Genomes Project Consortium [10], and (c) the variant impact should be either modifier, moderate,or high, according to SnpEff classification (https://pcingola.github.io/SnpEff/; accessed on 20 January 2022)—a program that predicts coding effects such as synonymous or non-synonymous amino acid replacement, start codon gains or losses, stop codon gains or losses, and/or frameshifts. Predicted effects concerned protein-coding genes [26]. A total of 64 variants were found in the *ABO*, *CCR9*, *CXCR6*, *FYCO1*, *LZTFL1*, *XCR1*, and *SLC6A20* genes, and are described in Supplementary Table S1. The analyses were directed to 38 variants that met all specifications of the selection criteria. The function of these genes is summarized in Table 1.

Table 1. Function of the *SLC6A20, LZTFL1*, *CCR9, CXCR6, XCR1, FYCO1,* and *ABO* genes.

Gene	Description *
SLC6A20	This gene encodes the protein sodium–amino acid(proline) transporter 1 (SIT1), which interacts with the angiotensin-converting enzyme 2 (*ACE2*)—the SARS-CoV-2 cell-surface receptor—allowing its heterodimerization [19]. The heterodimerization of the *ACE2* protein is necessary for the formation of a quaternary structure that functions as a binding site for the SARS-CoV-2 protein S [27].
LZTFL1	The *LZTFL1* gene encodes the leucine zipper transcription factor-like 1, and its function is relatedto tumor-suppressor action and negative regulation of the hedgehog signaling pathways. This gene has high expression in lung tissues [25,28]; it is related to the functioning of the cilia of the pulmonary epithelium and to the signaling of important intracellular pathways, regulating the epithelial–mesenchymal transformation [29].
CCR9	CC chemokines are mainly responsible for the recruitment of lymphocytes. *CCR9* is the receptor for the C-C chemokine ligand 25 (CCL25). The CCR9 receptor is mainly found on immature T lymphocytes and the surface of intestinal cells [30]. Animal studies have shown that the CCR9/CCL25 complex participates in the action of T helper 1 (Th1) cells. Another finding indicates that in knockout rats there was a reduction in the mRNA levels of pro-inflammatory cytokines (i.e., IL-6, IL-1β, and TNF-α) [30,31].
CXCR6	CXC chemokines have the highest ability to attract neutrophils and monocytes (30). *CXCR6* is the receptor for *CXCL16*; in cellular studies and animal models, it has been shown to regulate inflammatory activity and influence the levels of INF-γ and TNF-α secreted by CD4+ T cells [32,33].
XCR1	*XCR1* encodes the receptor of the XCL-1 ligand. The receptor triggers chemotactic signals in the presence ofthe ligand [34]. *XCR1* is expressedin the lung tissue. Further reports suggest that *XCL1* expression in NKcells and CD8+ T cells is constitutively detectable at a steady state, and is elevated during viral infection in mice and humans. The XCL1–XCR1 axis is important for efficient cytotoxic immune response mediated by CD8+ T cells [35].
FYCO1	This gene is responsible for the production of a Rab7 adapter protein, and has the function of assisting in the intracellular transport of autophagic vesicles via transport by microtubules. To carry out the transport, the encoded protein interacts with Rab7 GTPase, phosphatidylinositol-3-phosphate (PI3P), the autophagosome marker LC3, and the kinesin KIF5 [36,37]; it was previously found to be related to inclusion body myositis (IBM) and autosomal recessive congenital cataracts (CATC2) [38,39].
ABO	The *ABO* gene encodes the enzyme alpha 1-3-galactosyltransferase, which transforms the H antigen expressed on the cell surface of several cell types into A and B antigens. Furthermore, the enzyme convertsthe Hantigeninto the von Willebrand factor [40]. Studies indicate that group A confers risk of developing severe forms ofinfection, while group O confers protection[1]. This effect is related to the expressionof anti-A and anti-Bantibodies that could neutralize the interactionof the virus protein S with ACE2, blocking its adsorption [41]. Another hypothesis would be its actionin the formation of the von Willebrand factor and its relationship with its expression in the pulmonary endothelium, indirectly influencing pro-inflammatory regulation and cell adhesion [42,43].

* The gene functions related to COVID-19 are hypotheses raised by other authors.

3. Results

In our study, we identified 64 polymorphisms distributed in 7 genes: 14 of them from the *ABO* gene, 3 from the *CCR9* gene, 1 from the *CXCR6* gene, 1 from the *XCR1* gene, 11 from the *SLC6A20* gene, 4 from the *LZTFL1* gene, and 30 from the *FYCO1* gene (Supplementary Table S1). Among these 64 variants, only 47 were new variants or had impact prediction by SnpEff of high, moderate, or modifier. A total of 10 SNPs were located in the *ABO* gene, 2 in the *CCR9* gene, 1 in the *CXCR6* gene, 21 in the *FYCO1* gene, 4 in the *LZTFL1*gene, and 8 in the *SLC6A20* gene.

These variants are described in Table 2, which contains characteristics including SNP ID, chromosomal region, nucleotide change, SnpEff software impact prediction, and allele frequency for the NAM population and the five continental populations present in the 1000 Genomes platform (AFR, AMR, EAS, EUR, and SAS).

Table 2. Description of the variants with predicted high, moderate, or modifier impact, and new variants for the genes *ABO*, *CCR9*, *CXCR6*, *FYCO1*, *LZTFL1*, and *SLC6A20*.

Gene	Position	SNP ID	Ref [a]	Var [b]	Impact Predicted by SnpEff	Variant Allele Frequency					
						NAM	AFR	AMR	EAS	EUR	SAS
ABO	133256189	rs55727303	C	T	High	0.281	0.001	0.108	-	0.012	0.026
ABO	133255902	rs8176748	C	T	Moderate	0.814	0.243	0.431	0.290	0.221	0.220
ABO	133256085	rs8176740	A	T	Moderate	0.814	0.242	0.431	0.289	0.221	0.220
ABO	133256264	rs1053878	G	A	Moderate	0.016	0.256	0.068	0.151	0.099	0.029
ABO	133257465	rs8176721	G	A	Moderate	0.000	0.155	0.017	-	0.005	-
ABO	133257486	rs8176720	T	C	Moderate	0.814	0.489	0.494	0.483	0.337	0.464
ABO	133257246	rs2073824	A	G	Modifier	0.728	0.469	0.496	0.549	0.336	0.465
ABO	133257320	rs2073825	A	T	Modifier	0.235	0.241	0.431	0.289	0.222	0.225
ABO	133262062	*	C	A	Modifier	0.016	-	-	-	-	-
ABO	133275050	rs616154	C	T	Modifier	0.031	0.531	0.408	0.408	0.535	0.631
ABO	133275068	rs559723	A	G	Modifier	0.184	0.531	0.408	0.408	0.536	0.631
CCR9	45894830	rs7648467	C	A	Modifier	0.014	0.448	0.050	-	0.013	0.008
CCR9	45897524	rs17764980	G	A	Modifier	0.000	0.005	0.058	0.004	0.120	0.383
CXCR6	45946488	rs2234355	G	A	Moderate	0.033	0.491	0.068	-	0.005	0.001
FYCO1	45959401	*	G	A	Low	0.019	-	-	-	-	-
FYCO1	45923752	rs35678722	G	A	Moderate	0.083	0.012	0.001	-	-	-
FYCO1	45966331	rs13079478	G	T	Moderate	0.070	0.005	0.059	0.004	0.122	0.360
FYCO1	45966333	rs13059238	T	C	Moderate	0.070	0.019	0.063	0.004	0.123	0.359
FYCO1	45966722	rs113517878	C	T	Moderate	0.083	0.003	0.003	-	-	-
FYCO1	45967298	rs3796375	G	A	Moderate	0.822	0.093	0.565	0.661	0.431	0.372
FYCO1	45967995	rs33910087	G	A	Moderate	0.070	0.017	0.059	0.005	0.122	0.359
FYCO1	45968372	rs3733100	C	G	Moderate	0.885	0.210	0.643	0.667	0.556	0.731
FYCO1	45968585	rs4683158	C	T	Moderate	0.994	0.982	0.914	1.000	0.801	0.922
FYCO1	45979767	rs1306733846	C	T	Moderate	0.000	0.000	0.000	0.000	0.000	0.000
FYCO1	45923467	rs6800954	C	T	Modifier	0.143	0.287	0.193	0.307	0.216	0.148
FYCO1	45936407	rs1873002	T	C	Modifier	1.000	1.000	1.000	1.000	1.000	1.000
FYCO1	45938385	rs9875616	G	A	Modifier	0.859	0.914	0.850	0.954	0.746	0.879
FYCO1	45959378	rs13069079	G	A	Modifier	0.000	0.005	0.058	0.004	0.121	0.371
FYCO1	45959571	rs1532071	G	A	Modifier	0.908	0.260	0.614	0.652	0.529	0.732
FYCO1	45959735	rs76597151	G	A	Modifier	0.014	0.017	0.062	0.004	0.122	0.371
FYCO1	45969944	rs17214952	A	G	Modifier	0.014	0.019	0.063	0.004	0.123	0.360
FYCO1	45973053	rs41289622	T	G	Modifier	0.070	0.005	0.059	0.004	0.122	0.361
FYCO1	45975359	rs751552	A	T	Modifier	0.814	0.063	0.565	0.661	0.431	0.371
FYCO1	45981341	rs36122610	G	A	Modifier	0.054	0.005	0.059	0.004	0.122	0.358
FYCO1	45984767	rs3733097	G	A	Modifier	0.853	0.067	0.561	0.655	0.432	0.372
LZTFL1	45828480	rs1129183	C	T	Moderate	0.000	0.043	0.038	0.039	0.074	0.077
LZTFL1	45827235	*	TCTG	T	Modifier	0.016	-	-	-	-	-
LZTFL1	45842023	rs138230559	C	T	Modifier	0.009	0.033	-	-	-	-
LZTFL1	45842083	rs141398338	A	C	Modifier	0.083	0.005	-	-	0.002	-
SLC6A20	45759079	rs140440513	C	T	Moderate	0.083	0.000	0.000	0.000	0.000	0.000
SLC6A20	45759901	rs61731475	T	C	Moderate	0.000	-	0.006	-	0.014	-
SLC6A20	45772602	rs17279437	G	A	Moderate	0.017	0.005	0.043	0.006	0.092	0.031
SLC6A20	45775922	rs139429025	T	C	Moderate	0.000	0.012	-	-	-	-
SLC6A20	45758379	rs2251347	C	T	Modifier	0.994	0.990	0.976	0.921	0.954	0.972
SLC6A20	45760066	rs116638840	C	T	Modifier	0.027	0.076	0.010	-	-	0.003
SLC6A20	45762899	rs2191027	C	T	Modifier	0.014	0.020	0.193	0.017	0.299	0.149
SLC6A20	45780132	rs2252547	T	C	Modifier	0.155	0.465	0.614	0.450	0.591	0.516

[a] Reference allele; [b] variant allele; * variants without described SNP; (-) no annotation; NAM: Amazonian Native American populations; AFR: African populations; AMR: American populations; EAS: East Asian populations; EUR: European populations; SAS: South Asian populations. All variants described in this table follow the selection criteria based on the impact prediction of modifier, moderate, or high, according to SnpEff.

We identified three new variants in the Amazonian Native American population: One of the variants was located in the *ABO* gene at position 133262062, with base exchange C > A, in the intronic region, with an allele frequency of 0.016. The second variant was identified in the *FYCO1* gene at position 45959401, with base exchange G > A, in the exonic region, withlow impact predicted by SnpEff and an allele frequency of 0.018. The third variant was identified in the *LZTFL1* gene at position 45827235, with a TCTG > T deletion, in the intronicregion, and with an allele frequency of 0.016.

Among the polymorphisms analyzed, eight showed no variant allele frequency in Amerindian populations: two of these were in the *ABO* gene (rs200932155 and rs8176721), one in the *CCR9* gene (rs17764980), two in the *FYCO1* gene (rs3796376 and rs13069079), one in the *LZTFL1* gene (rs1129183), and two in the *SLC6A20* gene (rs61731475 and rs139429025).

Regarding the classification based on impact forecast by SnpEff, the polymorphism in the *ABO* gene (rs55727303) presented a high impact with a significant difference in all of the correlations. We identified 15 polymorphisms with moderate impact prediction in 4 genes, distributed in 4 polymorphisms in the *ABO* gene (rs8176748, rs8176740, rs8176720, and rs1053878), 1 in the *CXCR6* gene (rs2234355), 8 in the *FYCO1* gene (rs3796375, rs35678722, rs113517878, rs4683158, rs13079478, rs13059238, rs33910087, and rs3733097), and 2 in the *SLC6A20* gene (rs140440513 and rs17279437).

Additionally, 38 variants had allele frequency in all populations—a necessary requirement for comparative analysis between the populations studied using Fisher's exact test (Table 2). The remaining variants were excluded because they had no frequency description in the 1000 Genomes Project.

Among the variants analyzed, 18 showed significant differences of the NAM population when compared to all other continental populations (AFR, AMR, EAS, EUR, and SAS): 7 belonging to the *ABO* gene (rs55727303, rs8176748 rs8176740, rs8176720, rs2073824, rs559723, and rs616154), 7 from the *FYCO1* gene (rs3733100, rs3796375, rs35678722, rs113517878, rs1532071, rs3733097, and rs751552), 2 from the *SLC6A20* gene (rs2252547 and rs140440513), 1 from the *LZTFL1* gene (rs141398338), and 1 from the *CCR9* gene (rs147314165). The remaining 20 variants that were not significant in any comparisons are described in Table 3.

Figure 1 represents the genomic differences between the studied populations in multidimensional scaling (MDS), based on the fixation index (FST) of polymorphisms. We can observe the presence of a genetically similar core set composed of the AMR, EUR, and SAS components, while the other groups (NAM, EAS, and AFR) are at extreme points of the graph.

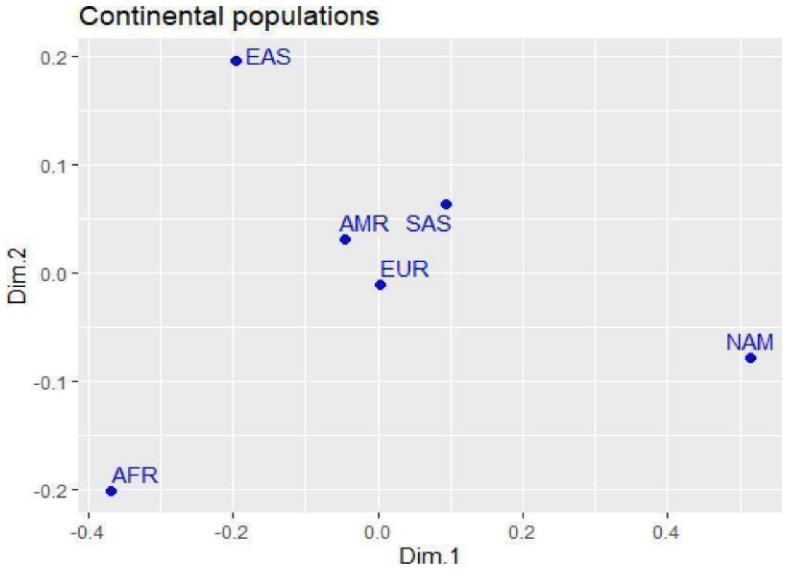


Figure 1. Differences in allele frequencies of the variants studied in the continental populations and the Native American population, plotted in MDS.

Table 3. Comparison of the allele frequencies between the NAM population and the continental populations (AFR, AMR, EUR, EAS, and SAS).

Gene	SNP ID	NAM vs. AFR *	NAM vs. AMR *	NAM vs. EAS *	NAM vs. EUR *	NAM vs. SAS *
ABO	rs55727303	**1.66×10^{-19}**	**1.08×10^{-3}**	**1.40×10^{-17}**	**6.75×10^{-14}**	**8.82×10^{-11}**
ABO	rs8176748	**1.58×10^{-19}**	**1.00×10^{-8}**	**7.06×10^{-16}**	**1.52×10^{-20}**	**1.29×10^{-20}**
ABO	rs8176740	**1.58×10^{-19}**	**1.00×10^{-8}**	**7.06×10^{-16}**	**1.52×10^{-20}**	**1.29×10^{-20}**
ABO	rs8176720	**3.78×10^{-7}**	**2.10×10^{-6}**	**4.54×10^{-7}**	**3.81×10^{-13}**	**9.84×10^{-8}**
ABO	rs1053878	**6.04×10^{-7}**	0.149	**0.001**	**0.020**	1.000
ABO	rs2073824	**6.48×10^{-5}**	**5.73×10^{-4}**	**4.73×10^{-3}**	**2.04×10^{-9}**	**5.21×10^{-5}**
ABO	rs559723	**7.94×10^{-8}**	**6.70×10^{-4}**	**5.66×10^{-4}**	**9.77×10^{-8}**	**1.11×10^{-11}**
ABO	rs616154	**7.78×10^{-17}**	**5.92×10^{-11}**	**5.50×10^{-11}**	**6.58×10^{-17}**	**3.03×10^{-22}**
ABO	rs2073825	1.000	**3.0×10^{-3}**	0.382	0.874	0.874
CCR9	rs147314165	**4.40×10^{-4}**	**4.22×10^{-4}**	**8.63×10^{-5}**	**8.70×10^{-5}**	**9.84×10^{-5}**
CCR9	rs7648467	**2.45×10^{-14}**	0.330	0.113	0.570	0.461
CXCR6	rs2234355	**5.37×10^{-15}**	0.398	**0.035**	0.100	**0.037**
FYCO1	rs3733100	**4.13×10^{-28}**	**3.67×10^{-5}**	**1.45×10^{-4}**	**5.81×10^{-8}**	**5.27×10^{-3}**
FYCO1	rs3796375	**2.35×10^{-37}**	**4.62×10^{-5}**	**6.62×10^{-3}**	**1.20×10^{-9}**	**3.18×10^{-12}**
FYCO1	rs35678722	**3.39×10^{-5}**	**4.22×10^{-4}**	**8.63×10^{-5}**	**8.70×10^{-5}**	**9.84×10^{-5}**
FYCO1	rs113517878	**8.44×10^{-5}**	**4.22×10^{-4}**	**8.63×10^{-5}**	**8.70×10^{-5}**	**9.84×10^{-5}**
FYCO1	rs4683158	0.614	0.008	1.000	**2.49×10^{-6}**	**0.015**
FYCO1	rs13079478	**4.40×10^{-4}**	0.578	**2.76×10^{-4}**	0.410	**1.51×10^{-6}**
FYCO1	rs13059238	**0.012**	0.591	**2.76×10^{-4}**	0.410	**1.51×10^{-6}**
FYCO1	rs33910087	**0.009**	0.578	**0.001**	0.410	**1.51×10^{-6}**
FYCO1	rs1532071	**6.14×10^{-25}**	**1.36×10^{-6}**	**1.32×10^{-5}**	**8.95×10^{-10}**	**1.82×10^{-3}**
FYCO1	rs3733097	**2.08×10^{-45}**	**2.84×10^{-6}**	**6.14×10^{-4}**	**2.84×10^{-11}**	**7.34×10^{-14}**
FYCO1	rs751552	**1.41×10^{-41}**	**1.51×10^{-4}**	**1.54×10^{-2}**	**7.09×10^{-9}**	**2.68×10^{-11}**
FYCO1	rs1873002	1.000	1.000	1.000	1.000	1.000
FYCO1	rs9875616	0.168	1.000	**6.0×10^{-3}**	**0.045**	0.685
FYCO1	rs6800954	**0.012**	0.383	**5.0×10^{-3}**	0.192	1.000
FYCO1	rs41289622	**4.40×10^{-4}**	0.578	**2.76×10^{-4}**	0.410	**1.51×10^{-6}**
FYCO1	rs36122610	**0.011**	1.000	**0.012**	0.093	**2.44×10^{-8}**
FYCO1	rs76597151	1.000	0.230	0.302	**5.0×10^{-3}**	**5.07×10^{-11}**
FYCO1	rs17214952	1.000	0.230	0.302	**5.0×10^{-3}**	**1.28×10^{-10}**
LZTFL1	rs141398338	**4.40×10^{-4}**	**4.22×10^{-4}**	**8.63×10^{-5}**	**8.70×10^{-5}**	**9.84×10^{-5}**
LZTFL1	rs138230559	0.712	0.288	0.213	0.213	0.218
SLC6A20	rs140440513	**2.59×10^{-5}**	**4.22×10^{-4}**	**8.63×10^{-5}**	**8.70×10^{-5}**	**9.84×10^{-5}**
SLC6A20	rs17279437	0.371	0.485	0.381	0.05	1.000
SLC6A20	rs2252547	**1.08×10^{-14}**	**6.54×10^{-12}**	**3.18×10^{-6}**	**2.12×10^{-11}**	**1.95×10^{-8}**
SLC6A20	rs2251347	1.000	0.366	**0.016**	0.095	0.380
SLC6A20	rs116638840	0.306	0.236	**0.035**	**0.035**	0.068
SLC6A20	rs2191027	1.000	**7.90×10^{-5}**	1.000	**3.58×10^{-8}**	**1.0×10^{-3}**

NAM: Amazonian Native American populations; AFR: African populations; AMR: American populations; EAS: East Asian populations; EUR: European populations; SAS: South Asian populations; * p-value defined by Fisher's exact test. Bold characters indicate a significant difference (p-value < 0.05).

4. Discussion

COVID-19 presents a new threat to the health of Native Amerindians living remotely.In the Amazon region, it is estimated that there are around 78 Amerindian tribes living in isolation [44]. The Amerindian people belong to a vulnerable population, who lack immunity to many infectious diseases [8]. In Brazilian territory, there were 59,574 cases and 871 deaths recorded in Amerindians as of 6 February 2022 [16].

The occurrence of numerous stochastic events—such as geographic isolation, inbreeding, and genetic drift—may contribute to genetic differentiation in Amazonian Native American populations [45–49]. These variables influence the formation of indigenous peoples, their ethnic structure and, consequently, their genomic patterns, with different allele frequencies from other continental populations [50,51].

The genomic differences and distinct sociodemographic and anthropological characteristics of Amerindian populations are related to epidemiological differences in respiratory and viral diseases—such as tuberculosis (TB), human immunodeficiency virus (HIV), hu- man T-lymphotropic virus (HTLV), and human herpesvirus type 8 (HHV-8)—observed in comparisons between Amerindian and non-Amerindian populations. Studies suggest

that genomic alterations in the immune response patterns and the parasite–host molecular interaction may be the causes of the observed differences [7,17,52].

GWA studies are widely used in the identification of genetic variants associated with complex and multifactorial diseases, such as cardiovascular, psychiatric, infectious, and numerous other diseases [53–55]. The first GWA study investigated the multigenic groupof chromosome 3 in patients infected with COVID-19 [19]. The 3p21.31 locus has the *XCR1*, *CCR9*, *CXCR6*, *SLC6A20*, *LZTFL1*, and *FYCO1* genes, in which rs11385942 (*LZTFL1*) and rs657152 (*ABO*) were significantly associated with severe forms of the disease [19].

In Shelton's 2021 study, the variants related to disease severity and susceptibility were rs13078854 of the *LZTFL1* gene and rs9411378 of the *ABO* gene [20]. These findings reinforce the importance of the 3p21.31 locus and the 9q34.21 locus—particularly with regard to the *ABO* and *LZTFL1* genes; therefore, we focus our discussion on these two genes.

The association between the *ABO* blood group and COVID-19 infection and severity was studied. Blood type A might be more susceptible to COVID-19 infection, while blood type O might be less susceptible to this disease. [19]. In other epidemiological and genomic studies, susceptibility was also lower in O blood group patients [20,56,57]. The O allele is the most frequent blood type found in the Native American population [51,58]. In addition, the *ABO* blood group has previously been linked to susceptibility to other diseases, such as influenza, malaria, schistosomiasis, and SARS-CoV-2 [59].

Some proposed mechanisms for the association between ABO blood type and SARS-CoV-2 infection were investigated, as follows: (1) anti-A and/or anti-B antibodies play a role as viral neutralizing antibodies when binding to A and/or B antigens expressed on the viral envelope; (2) the SARS-CoV-2 S protein is bound by human anti-A antibodies, and prevents entry into the lung epithelium when blocking the interaction between the virus and ACE2R; (3) an increase in ACE-1 activity in group A individuals can cause predisposition to cardiovascular disease and lead to severe COVID-19; (4) ABH glycans in the SARS-CoV-2 S protein may modify cellular receptors of SARS-CoV-2 for ACE2R;
(5) ABH glycans on target cells could serve as alternative, lower affinity receptors for the SARS-CoV-2 S protein, or could bind other viral envelope structures [42,60,61].

In our study, significant differences were found in the majority of polymorphisms in the *ABO* gene between the Amerindian populations and continental populations. We hypothesize that the differences in genomic profiles and the novel variants identified in the Native American population may influence the development of severe forms of COVID-19. However, further studies will be needed in COVID-19-positive individuals in this population in order to better understand the potential influence of these variants on this infection.

The *LZTFL1* gene encodes the leucine zipper transcription factor-like 1, and its function is related to tumor-suppressor action and negative regulation of the hedgehog signaling pathways. In knockout zebrafish (*Danio rerio*) experimental models, impaired cell traffic in ciliary membranes, retinal degeneration, and obesity were observed [28]. In addition, this gene has high expression in lung tissues; however, the mechanisms directly related to SARS-CoV-2 infection remain unknown [25].

In our study, we found four variants related to the *LZTFL1* gene, and only rs141398338 showed a significant difference in the NAM population when compared to the continental populations. There have been no reports in the literature on the association between rs141398338 and severe forms of COVID-19. New variants identified in the Native American population may influence the development of severe forms of COVID-19, and further human genetic studies need to be carried out in order to clarify this issue.

In addition, we identified three new variants in the Amazonian NAM population; these SNPs were located in the *FYCO1*, *ABO*, and *LZTFL1* genes. The first has low clinical impact, while the following two have modifier impact. These mutations—especially the oneswith modifier impact—could have important potential as markers of severe forms of COVID-19 in Amazonian indigenous populations, as well as intronic regional mutations

that significantly influence gene expression levels [62]. Larger studies should be performed to confirm these new variants in patients diagnosed with COVID-19.

Genetic variants associated with severe COVID-19 indicated by the studies of Elling- haus and Shelton et al. [19,20] were not found in the Amazonian Native American popula- tion. The allelic frequencies of the SNPs in the NAM group were lower than for any of the other groups in our study, showing that the Amazonian Native Americans have low genetic variability anda different genetic pool. Genetic variants present in the NAM population and low genetic variability could indicate a protective factor against severe COVID-19.

The limitation of our study was the small number of NAM individuals, who come from isolated and relatively small populations in the Amazonian region. This study is a preliminary severe COVID-19 study, and did not investigate individuals with COVID-19 in- fection. We collected blood samples from individuals before the COVID-19 pandemic. Our results may reveal important information and contribute to the assessment of individual risk for the development of this disease.

5. Conclusions

Genetic variants associated with severe COVID-19 were not found in the Amazonian Native American population. The allele frequency for the candidate genes in the NAM group was significantly different from the frequencies observed in continental groups. This may provide a protective factor against severe COVID-19. We also identified two new genetic variants with modifier impact in the Amazonian population that could be studied in order to validate the possible associations with COVID-19 outcomes. This work contributes to the elucidation of the genomic profile of Amazonian Native Americans—an understudied population—by providing genomic data that may help forthcoming studies to improve COVID-19 outcomes. Future studies should be performed in this population to identify more genetic variants related to severe COVID-19.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/jpm12040554/s1: Supplementary Table S1: Description of all variants for the genes *ABO*, *CCR9*, *CXCR6*, *FYCO1*, *LZTFL1*, *SLC6A20*, and *XCR1* found in the 64 individuals sampled in the present study.

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

Funding: This research was funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (http://www.cnpq.br), the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (https://www.gov.br/capes/pt-br), the Pró-Reitoria de Pesquisa e Pós-Graduação da UFPA (PROPESP) (http://www. propesp.ufpa.br), and the Fundação Amazônia de Amparo a Estudos e Pesquisas (FAPESPA) (http://www.fapespa.pa.gov.br). This work is part of the Rede de Pesquisa em Genômica Populacional Humana (Biocomputacional-Protocol no. 3381/2013/CAPES).

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the National Committee for Ethics in Research (CONEP) and the Research Ethics Committee of the UFPA Tropical Medicine Center, under CAAE number 20654313.6.0000.5172.

Informed Consent Statement: All participants in the study and their ethnic group leaders signed written informed consent. Written informed consent was obtained from the patients to publish this paper.

Data Availability Statement: The data presented in this study are openly available on Figshare at https://doi.org/10.6084/m9.figshare.18728192.v1; accessed on 25 January 2022.

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

References

1. Wang, W.; Tang, J.; Wei, F. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. *J. Med. Virol.* **2020**, *92*, 441–447. [CrossRef]
2. Cucinotta, D.; Vanelli, M. WHO Declares COVID-19 a Pandemic. *Acta Bio-Med. Atenei Parm.* **2020**, *91*, 157–160.
3. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. Available online: https://covid19.who.int (accessed on 6 February 2022).
4. Samadizadeh, S.; Masoudi, M.; Rastegar, M.; Salimi, V.; Shahbaz, M.B.; Tahamtan, A. COVID-19: Why does disease severity vary among individuals? *Respir. Med.* **2021**, *180*, 106356. [CrossRef] [PubMed]
5. Anastassopoulou, C.; Gkizarioti, Z.; Patrinos, G.P.; Tsakris, A. Human genetic factors associated with susceptibility to SARS-CoV-2 infection and COVID-19 disease severity. *Hum. Genom.* **2020**, *14*, 40. [CrossRef] [PubMed]
6. Mario-Vásquez, J.E.; Naranjo-González, C.A.; Montiel, J.; Zuluaga, L.M.; Vásquez, A.M.; Tobón-Castaño, A.; Bedoya, G.; Segura, C. Association of variants in IL1B, TLR9, TREM1, IL10RA, and CD3G and Native American ancestry on malaria susceptibility in Colombian populations. *Infect. Genet. Evol.* **2021**, *87*, 104675. [CrossRef]
7. Leal, D.F.D.V.B.; Da Silva, M.N.S.; Fernandes, D.C.R.D.O.; Rodrigues, J.C.G.; Barros, M.C.D.C.; Pinto, P.D.D.C.; Pastana, L.F.; Da Silva, C.A.; Fernandes, M.R.; De Assumpção, P.P.; et al. Amerindian genetic ancestry as a risk factor for tuberculosis in an amazonian population. *PLoS ONE* **2020**, *15*, e0236033. [CrossRef] [PubMed]
8. Castro e Silva, M.A.; Ferraz, T.; Couto-Silva, C.M.; Lemes, R.B.; Nunes, K.; Comas, D.; Hünemeier, T. Population Histories and Genomic Diversity of South American Natives. *Mol. Biol. Evol.* **2021**, *39*, msab339. [CrossRef] [PubMed]
9. Gilmour, B.; Alene, K.A.; Clements, A. The prevalence of tuberculosis and malaria in minority indigenous populations of South-East Asia and the Western Pacific Region: A systematic review and meta-analysis. *Pathog. Glob. Health* **2021**, 1–19. [CrossRef]
10. Pinto, P.; Salgado, C.; Santos, N.P.C.; Santos, S.; Ribeiro-dos-Santos, Â. Influence of Genetic Ancestry on INDEL Markers of NFKβ1, CASP8, PAR1, IL4 and CYP19A1 Genes in Leprosy Patients. *PLoS Negl. Trop. Dis.* **2015**, *9*, e0004050. [CrossRef]
11. Tai, D.B.G.; Shah, A.; Doubeni, C.A.; Sia, I.G.; Wieland, M.L. The Disproportionate Impact of COVID-19 on Racial and Ethnic Minorities in the United States. *Clin. Infect. Dis.* **2021**, *72*, 703–706. [CrossRef]
12. Da Cunha, A.A.; Corona, R.A.; Castilho-Martins, E.A. COVID-19 and race/color disparity: A brief analysis of the indigenous population in a state in the Brazilian Amazon. *Einstein* **2021**, *19*, eCE6734. [CrossRef]
13. Stone, M.J.; Close, R.M.; Jentoft, C.K.; Pocock, K.; Lee-Gatewood, G.; Grow, B.I.; Parker, K.H.; Twarkins, A.; Nashio, J.T.; McAuley, J.B. High-Risk Outreach for COVID-19 Mortality Reduction in an Indigenous Community. *Am. J. Public Health* **2021**, *111*, 1939–1941. [CrossRef]
14. Serrano-Coll, H.; Miller, H.; Rodríguez-Van, D.H.C.; Gastelbondo, B.; Novoa, W.; Oviedo, M.; Rivero, R.; Garay, E.; Mattar, S. High Prevalence of SARS-CoV-2 in an Indigenous Community of the Colombian Amazon Region. *Trop. Med. Infect. Dis.* **2021**, *6*, 191. [CrossRef]
15. Cupertino, G.A.; Cupertino, M.D.C.; Gomes, A.P.; Braga, L.M.; Siqueira-Batista, R. COVID-19 and Brazilian Indigenous Populations. *Am. J. Trop. Med. Hyg.* **2020**, *103*, 609–612. [CrossRef]
16. Secretaria de Saúde Indígena (SESAI). Boletim Epidemiológico 492, 2022, 1. Available online: https://www.gov.br/saude/pt-br/composicao/sesai (accessed on 30 November 2020).
17. Lindenau, J.D.; Guimarães, L.S.P.; Friedrich, D.C.; Hurtado, A.M.; Hill, K.R.; Salzano, F.M.; Hutz, M.H. Cytokine gene polymor- phisms are associated with susceptibility to tuberculosis in an Amerindian population. *Int. J. Tuberc. Lung Dis.* **2014**, *18*, 952–957. [CrossRef]
18. Wang, S.; Lewis, C.M.; Jakobsson, M.; Ramachandran, S.; Ray, N.; Bedoya, G.; Rojas, W.; Parra-Marín, M.V.; A Molina, J.; Gallo, C.; et al. Genetic Variation and Population Structure in Native Americans. *PLoS Genet.* **2007**, *3*, e185. [CrossRef]
19. Ellinghaus, D. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *N. Engl. J. Med.* **2020**, *383*, 1522–1534.
20. Shelton, J.F.; Shastri, A.J.; Ye, C.; Weldon, C.H.; Filshtein-Sonmez, T.; Coker, D.; Symons, A.; Esparza-Gordillo, J.; Aslibekyan, S.; Auton, A.; et al. Trans-ancestry analysis reveals genetic and nongenetic associations with COVID-19 susceptibility and severity. *Nat. Genet.* **2021**, *53*, 801–808. [CrossRef]
21. De Ramos, B.R.A.; D'Elia, M.P.B.; Amador, M.A.T.; Santos, N.P.C.; Santos, S.E.B.; da Castelli, E.C.; Witkin, S.S.; Miot, H.A.; Miot, L.D.B.; da Silva, M.G. Neither self-reported ethnicity nor declared family origin are reliable indicators of genomic ancestry. *Genetica* **2016**, *144*, 259–265. [CrossRef]
22. Rodrigues, J.C.G.; de Souza, T.P.; Pastana, L.F.; Ribeiro dos Santos, A.M.; Fernandes, M.R.; Pinto, P.; Wanderley, A.V.; De Souza, S.J.; Kroll, J.E.; Pereira, A.L.; et al. Identification of NUDT15 gene variants in Amazonian Amerindians and admixed individuals from northern Brazil. *PLoS ONE* **2020**, *15*, e0231651. [CrossRef]
23. Green, M.R.; Sambrook, J. Isolation of High-Molecular-Weight DNA Using Organic Solvents. *Cold Spring Harb. Protoc.* **2017**, pdb.prot093450. [CrossRef] [PubMed]
24. Ribeiro-dos-Santos, A.M.; Vidal, A.F.; Vinasco-Sandoval, T.; Guerreiro, J.; Santos, S.; Ribeiro-dos-Santos, Â. Exome Sequencing of Native Populations from the Amazon Reveals Patterns on the Peopling of South America. *Front. Genet.* **2020**, *11*, 548507. [CrossRef] [PubMed]

25. Elhabyan, A.; Elyaacoub, S.; Sanad, E.; Abukhadra, A.; Elhabyan, A.; Dinu, V. The role of host genetics in susceptibility to severeviral infections in humans and insights into host genetics of severe COVID-19: A systematic review. *Virus Res.* **2020**, *289*, 198163. [CrossRef]
26. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **2012**, *6*, 80–92. [CrossRef] [PubMed]
27. Camargo, S.M.R.; Vuille-dit-Bille, R.N.; Meier, C.F.; Verrey, F. ACE2 and gut amino acid transport. *Clin. Sci.* **2020**, *134*, 2823–2833. [CrossRef]
28. Wei, Q.; Chen, Y.; Gu, Y.F.; Zhao, W. Molecular Characterization and Functional Analysis of Leucine Zipper Transcription Factor like 1 in Zebrafish (*Danio rerio*). *Front. Physiol.* **2019**, *10*, 801. [CrossRef]
29. Downes, D.J.; Cross, A.R.; Hua, P.; Roberts, N.; Schwessinger, R.; Cutler, A.J.; Munis, A.M.; Brown, J.; Mielczarek, O.; de Andrea, C.E.; et al. Identification of LZTFL1 as a candidate effector gene at a COVID-19 risk locus. *Nat. Genet.* **2021**, *53*, 1606–1615. [CrossRef]
30. Wu, X.; Sun, M.; Yang, Z.; Lu, C.; Wang, Q.; Wang, H.; Deng, C.; Liu, Y.; Yang, Y. The Roles of CCR9/CCL25 in Inflammation and Inflammation-Associated Diseases. *Front. Cell Dev. Biol.* **2021**, *9*, 2244. [CrossRef]
31. Pathak, M.; Lal, G. The Regulatory Function of CCR9+ Dendritic Cells in Inflammation and Autoimmunity. *Front. Imunol.* **2020**, 2219.
32. Liu, G.; Abas, O.; Strickland, A.B.; Chen, Y.; Shi, M. CXCR6^{+}CD4^{+} T cells promote mortality during *Trypanosoma brucei* infection. *PLoS Pathog.* **2021**, *17*, e1009968. [CrossRef]
33. Luoma, A.M.; Suo, S.; Williams, H.L.; Sharova, T.; Sullivan, K.; Manos, M.; Bowling, P.; Hodi, F.S.; Rahma, O.; Sullivan, R.J.; et al. Molecular Pathways of Colon Inflammation Induced by Cancer Immunotherapy. *Cell* **2020**, *182*, 655–671.e22. [CrossRef]
34. Balan, S.; Saxena, M.; Bhardwaj, N. Chapter One—Dendritic cell subsets and locations. In *International Review of Cell and Molecular Biology*; Lhuillier, C., Galluzzi, L., Eds.; Academic Press: Cambridge, MA, USA, 2019; pp. 1–68. Available online: https://www.sciencedirect.com/science/article/pii/S193764481930067X (accessed on 30 November 2020).
35. Lei, Y.; Takahama, Y. XCL1 and XCR1 in the immune system. *Microbes Infect.* **2012**, *14*, 262–267. [CrossRef]
36. Pankiv, S.; Alemu, E.A.; Brech, A.; Bruun, J.-A.; Lamark, T.; Overvatn, A.; Bjørkøy, G.; Johansen, T. FYCO1 is a Rab7 effector that binds to LC3 and PI3P to mediate microtubule plus end-directed vesicle transport. *J. Cell Biol.* **2010**, *188*, 253–269. [CrossRef]
37. Kuhn, C.; Menke, M.; Senger, F.; Mack, C.; Dierck, F.; Hille, S.; Schmidt, I.; Brunke, G.; Brünger, P.; Schmiedel, N.; et al. FYCO1 Regulates Cardiomyocyte Autophagy and Pre-vents Heart Failure Due to Pressure Overload In Vivo. *JACC Basic Transl. Sci.* **2021**, *6*, 365–380. [CrossRef]
38. Iqbal, H.; Khan, S.Y.; Zhou, L.; Irum, B.; Ali, M.; Ahmed, M.R.; Shahzad, M.; Ali, M.H.; Naeem, M.A.; Riazuddin, S.; et al. Mutations in FYCO1 identified in families with congenital cataracts. *Mol. Vis.* **2020**, *26*, 334–344.
39. Rothwell, S.; Lilleker, J.B.; Lamb, J.A. Genetics in inclusion body myositis. *Curr. Opin. Rheumatol.* **2017**, *29*, 639. [CrossRef]
40. Franchini, M.; Crestani, S.; Frattini, F.; Sissa, C.; Bonfanti, C. ABO blood group and von Willebrand factor: Biological implications. *Clin. Chem. Lab. Med. (CCLM)* **2014**, *52*, 1273–1276. [CrossRef]
41. Breiman, A.; Ruvën-Clouet, N.; Le Pendu, J. Harnessing the natural anti-glycan immune response to limit the transmission of enveloped viruses such as SARS-CoV-2. *PLoS Pathog.* **2020**, *16*, e1008556. [CrossRef]
42. Goel, R.; Bloch, E.M.; Pirenne, F.; Al-Riyami, A.Z.; Crowe, E.; Dau, L.; Land, K.; Townsend, M.; Jecko, T.; Rahimi-Levene, N.; et al. ABO blood group and COVID-19: A review on behalf of the ISBT COVID-19 working group. *Vox Sang.* **2021**, *116*, 849–861. [CrossRef]
43. Murray, G.P.; Post, S.R.; Post, G.R. ABO blood group is a determinant of von Willebrand factor protein levels in human pulmonary endothelial cells. *J. Clin. Pathol.* **2020**, *73*, 347–349. [CrossRef]
44. IBGE. Censo Brasileiro de 2010. Rio de Janeiro-RJ. 2012. Available online: https://biblioteca.ibge.gov.br/visualizacao/periodicos/552/cd_2010_agsn_if.pdf (accessed on 30 November 2020).
45. Salzano, F.M. Fatores dete rminísticos e estocásticos no processo microevolucionário humano. *Actas V Congr. Latinoam. Genet.* **1982**, *1*, 81–89.
46. Amos, W.; Hoffman, J.I. Evidence that two main bottleneck events shaped modern human genetic diversity. *Proc. Soc. Biol. Sci.* **2010**, *277*, 131–137. [CrossRef] [PubMed]
47. Dobbin, E.A.F.; Medeiros, J.A.G.; Costa, M.S.C.R.; Rodrigues, J.C.G.; Guerreiro, J.F.; Kroll, J.E.; Souza, S.; de Assumpção, P.; Ribeiro-Dos-Santos, Â.; Santos, S.; et al. Identification of Variants (rs11571707, rs144848, and rs11571769) in the BRCA2 Gene Associated with Hereditary Breast Cancer in Indigenous Populations of the Brazilian Amazon. *Genes* **2021**, *12*, 142. [CrossRef] [PubMed]
48. O'Fallon, B.D.; Fehren-Schmitz, L. Native Americans experienced a strong population bottleneck coincident with European contact. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20444–20448. [CrossRef] [PubMed]
49. De Brito Vargas, L.; Beltrame, M.H.; Ho, B.; Marin, W.M.; Dandekar, R.; Montero-Martín, G.; Fernández-Viña, M.A.; Hurtado, A.M.; Hill, K.R.; Tsuneto, L.T.; et al. Remarkably Low KIR and HLA Diversity in Amerindians Reveals Signatures of Strong Purifying Selection Shaping the Centromeric KIR Region. *Mol. Biol. Evol.* **2022**, *39*, msab298. [CrossRef]

50. Battilana, J.; Fagundes, N.J.R.; Heller, A.H.; Goldani, A.; Freitas, L.B.; Tarazona-Santos, E.; Munkhbat, B.; Munkhtuvshin, N.; Krylov, M.; Benevolenskaia, L.; et al. Alu insertion polymorphisms in Native Americans and related Asian populations. *Ann. Hum. Biol.* **2006**, *33*, 142–160. [CrossRef]
51. Estrada-Mena, B.; Estrada, F.J.; Ulloa-Arvizu, R.; Guido, M.; Méndez, R.; Coral, R.; Canto, T.; Granados, J.; Rubí-Castellanos, R.; Rangel-Villalobos, H.; et al. Blood group O alleles in Native Americans: Implications in the peopling of the Americas. *Am. J. Phys. Anthropol.* **2010**, *142*, 85–94. [CrossRef]
52. Salzano, F.M. *The Amerindian Microcosm: Anthropology, Comparative History, Ecology, Genetics and Evolution*; Hutz, M.H., Bortolini, M.C., Eds.; Cambridge Scholars Publishing: Newcastle upon Tyne, UK, 2019; p. 607.
53. Hart, A.B.; Kranzler, H.R. Alcohol Dependence Genetics: Lessons Learned from Genome-Wide Association Studies (GWAS) and Post-GWAS Analyses. *Alcohol. Clin. Exp. Res.* **2015**, *39*, 1312–1327. [CrossRef]
54. Nikpay, M.; Goel, A.; Won, H.-H.; Hall, L.M.; Willenborg, C.; Kanoni, S.; Saleheen, D.; Kyriakou, T.; Nelson, C.P.; Hopewell, J.C.; et al. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.* **2015**, *47*, 1121–1130.
55. Ko, D.C.; Urban, T.J. Understanding Human Variation in Infectious Disease Susceptibility through Clinical and Cellular GWAS. *PLoS Pathog.* **2013**, *9*, e1003424. [CrossRef]
56. Chen, T.; Wu, D.; Chen, H.; Yan, W.; Yang, D.; Chen, G.; Ma, K.; Xu, D.; Yu, H.; Wang, H.; et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: Retrospective study. *BMJ* **2020**, *368*, m1091. [CrossRef]
57. Zhao, J.; Yang, Y.; Huang, H.; Li, D.; Gu, D.; Lu, X.; Zhang, Z.; Liu, L.; Liu, T.; Liu, Y.; et al. Relationship between the ABO Blood Group and the Coronavirus Disease 2019 (COVID-19) Susceptibility. *Clin. Infect. Dis.* **2021**, *73*, 328–331. [CrossRef]
58. Halverson, M.S.; Bolnick, D.A. An ancient DNA test of a founder effect in native American ABO blood group frequencies. *American J. Phys. Anthropol.* **2008**, *137*, 342–347. [CrossRef]
59. Fricke-Galindo, I.; Falfán-Valencia, R. Genetics Insight for COVID-19 Susceptibility and Severity: A Review. *Front. Immunol.* **2021**, *12*, 1057. [CrossRef]
60. Cooling, L. Blood Groups in Infection and Host Susceptibility. *Clin. Microbiol. Rev.* **2015**, *28*, 801–870. [CrossRef]
61. Walls, A.C.; Park, Y.-J.; Tortorici, M.A.; Wall, A.; McGuire, A.T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181*, 281–292.e6. [CrossRef]
62. Rigau, M.; Juan, D.; Valencia, A.; Rico, D. Intronic CNVs and gene expression variation in human populations. *PLoS Genet.* **2019**, *15*, e1007902. [CrossRef]

3. ARTIGO 2: Perfil Molecular de Variantes Potencialmente Associadas a Formas Graves de COVID-19 em Populações Indígenas da Amazônia

Article

Molecular Profile of Variants Potentially Associated with Severe Forms of COVID-19 in Amazonian Indigenous Populations

Rita de Cássia Calderaro Coelho[1]†, Carlliane Lima e Lins Pinto Martins[1]†, Lucas Favacho Pastana[1], Juliana C. G. Rodrigues[1], Kaio Evandro Cardoso Aguiar[1], Amanda de Nazaré Cohen-Paes[1], Laura Patrícia Albarello Gellen[1], Francisco Cezar Aquino de Moraes[1*], Maria Clara Leite Calderaro[1], Letícia Almeida de Assunção[1], Natasha Monte[1], Esdras E. B. Pereira[1],André Maurício Ribeiro-dos-Santos[2], Andrea Campos Ribeiro dos Santos[2], Rommel Mario Rodriguez Burbano[1,3], Sandro José de Souza[4], João Farias Guerreiro[2], Paulo Pimentel de Assumpção[1], Sidney Emanuel Batista dos Santos[1,2], Marianne Rodrigues Fernandes[1,3], Ney Pereira Carneiro dos Santos[1,2]

[1] Laboratório do Núcleo de Pesquisa em Oncologia, Universidade Federal do Pará, Belém 66073-000, Brazil; rccalderarocoelho@gmail.com (R.d.C.C.C.) carllianepmartins@hotmail.com (C.L.L.P.M); lucas.pastana@ics.ufpa.br (L.F.P.); julianacgrodrigues@gmail.com (J.C.G.R.); kaio.evandro41@gmail.com (K.E.C.); laura.patricia.agellen@hotmail.com (L.P.A.G.); francisco.cezar2205@gmail.com (F.C.A.M); maria.calderaro@ics.ufpa.br (M.C.L.C.); leticiaalmeidaenf96@gmail.com (L.A.d.A.); ntshmonte@gmail.com (N.M.); edgarpereira@ufpa.br (E.E.B.P.); assumpcaopp@gmail.com (P.P.A.); sidneysantos@ufpa.br (S.E.B.S.); fernandesmr@yahoo.com.br (M.R.F.); npcsantos.ufpa@gmail.com (N.P.C.d.S.)

[2] Laboratório de Genética Humana e Médica, Universidade Federal do Pará, Belém 66075-110, Brazil; andremrsantos@gmail.com (A.M.R.-d.-S.); akelyufpa@gmail.com (Â.C.R.-d.-S.); joao.guerreiro53@gmail.com (J.F.G.);

[3] Hospital Ophir Loyola, Secretaria de Estado da Saúde do Pará, 66063-240, Brazil; rommelburbano@gmail.com (R.M.R.B.)

[4] Instituto do Cérebro, Universidade Federal do Rio Grande do Norte, 59076-550, Brazil; sandro@neuro.ufrn.br (S.J.d.S.);

* Correspondence: francisco.cezar2205@gmail.com

† These authors also contributed

Coronavirus disease 2019 (COVID-19) is an infection caused by SARS-CoV-2. GWAS studies suggest a strong association of genetic factors with the severity of the disease. However, many of these works are developed in European populations, and little is known about the genetic variability of indigenous peoples underlying infection by SARS-CoV-2. The objective of the study is to investigate genetic variants present in the *AQP3, ARHGAP27, ELF5L IFNAR2, LIMD1, OAS1* and *UPK1A* genes selected to the association with the severity of COVID-19 in a sample of indigenous people from the Brazilian Amazon. We performed the complete sequencing of the exome of 64 healthy indigenous people from the Brazilian Amazon. The allele frequency data of the population were compared with data from other continental populations. 66 variants present in the 7 genes studied due were identified, including a variant with high impact on the *ARHGAP27* gene (rs201721078) and 3 new variants located in the INDG population present in the *AQP3, IFNAR2 and LIMD1* genes, with low, moderate and modifying impact, respectively.

1. Introduction

In December 2019, in Wuhan, China, the first cases of severe acute respiratory infection (SARS-CoV-2) by coronavirus (COVID-2) were diagnosed [1-2]. In a short time, the cases of infection reached large proportions, spreading to several regions and reaching several countries [3-4]. According to data from the World Health Organization (WHO), on

June 7, 2023, 767,750,853 cases of SARS-CoV-2 infection and 6,941,095 COVID-19 deaths were confirmed worldwide [5]. Showing that, even with the improvement of the reality of the pandemic period, due to the administration of vaccine doses, this disease still represents a global public health problem [6].

SARS-CoV-2 infection can happen to anyone. However, there is a great clinical diversity among patients with COVID-19. The disease can manifest itself in different ways, varying from asymptomatic, mild and severe forms, and can lead to death [3-4]. Clinical studies relate the heterogeneity of the disease with the genetic influence on the individual response to infection. The association of genetic factors with the severity and clinical evolution of COVID-19 is still little investigated. However, genome-wide association studies (GWAS) have been developed to better understand the relationship of genes associated with the severity of this disease [7-8-9], enabling the development of more specialized therapies for the risk group [7].

The association of genetic factors with the severity of COVID-19 has been addressed in different studies [7-8-9]. Another study developed by our research group [10] also investigated the interaction of genes (*SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6, XR1* and *ABO*) with the most severe forms of the disease and demonstrate a strong relationship of locus 3p21.31 with the severity of SARS-CoV-2 infection [7-8].

Seven new genes (*AQP3, ARHGAP27, ELF5, IFNAR2, LIMD1, OAS1* and *UPK1A*) were related to the severity of COVID-19 in populations of European origin. The difference in disease severity between genders and the association of androgenic hormones with the severity of SARS-CoV-2 infection were also addressed in the study [9].

Therefore, the objective of our study is to investigate genetic variants in the genes AQP3, ARHGAP27, ELF5, IFNAR2, LIMD1, OAS1 and UPK1A that were selected due to the association of these genes with the severity of COVID-19, in a sample of Amazonian indigenous peoples.

2. Results

In our study, we identified 66 variants distributed in genes: of them present in the 7 *AQP3* gene, 14 in the *ARHGAP27,* 6 in *ELF5,* 14 in the *IFNAR2* gene, 11 in the *LIMD1* gene, 9 in *OAS1* and in the 5 *UPK1A* gene (Supplementary Table S1). 18 variants were excluded due to the low coverage. After going through these quality criteria and impact prediction, we identified three unique variants of the indigenous population (Table 2) and 45 more variants were included in the study (Table 1) and compared with the other world populations present on the 1000 Genomes platform (AFR, AMR, EAS, EUR and SAS).

Table 1 – Description of the variants according to the high, moderate and modifier impact present in the *AQP3, ARHGAP27, ELF5, IFNAR2, LIMD1, OAS1* and *UPK1A* genes.

Gene	Position	SNP ID	Ref[a]	Var[b]	Impact Predicted by SNPeff	Variant Allele Frequence					
						INDG	AFR	AMR	EAS	EUR	SAS
ARHGAP27	45404053	rs201721078	G	A	HIGH	0,218	0.0009	0.0421	-	0.0001	-
ARHGAP27	45429642	rs2959953	G	C	MODERATE	0,375	0.6953	0.6537	0.520	0.717	0.673
ARHGAP27	45429931	rs12949256	C	T	MODERATE	0	-	-	-	0.5	0.166
ARHGAP27	45395459	rs117139057	G	C	MODERATE	0,125	0.089	0.2028	0.2768	0.1791	0.1189
ARHGAP27	45430283	rs7222444	A	G	MODIFIER	0,009	0.102	0.158	0.0005	0.243	0.071
ARHGAP27	45395941	rs62064597	C	G	MODIFIER	0,083	0.0424	0.132	0.0008	0.2141	0.073
ARHGAP27	45429583	rs115993362	C	G	MODIFIER	0,083	0.147	00.0062	-	.0002	0.0006

ARHGAP27	45410208	rs35327136	C	A	MODIFIER	0,015	0.1308	0.1778	-	0.2233	0.0682
ARHGAP27	45396860	rs7213200	G	A	MODIFIER	0,016	-	-	-	-	-
ARHGAP27	45410198	rs142163608	C	T	MODIFIER	0,083	0.0162	-	-	-	-
ARHGAP27	45404234	rs184103721	C	G	MODIFIER	0,045	0.0004	0.048	0.010	0.0002	0.0004
ARHGAP27	45410058	rs35389313	T	C	MODIFIER	0,013	-	-	-	-	-
AQP3	33442274	rs2231235	G	A	MODIFIER	0	0.012	0.109	0.063	0.069	0.085
AQP3	33442988	rs2231231	A	C	MODIFIER	0,937	0.775	0.7272	0.721	0.639	0.586
AQP3	33447549	rs12555686	G	T	MODIFIER	0	0.0221	0.255	0.175	0.095	0.095
IFNAR2	33241950	rs1051393	T	G	MODERATE	0,789	0.193	0.487	0.583	0.336	0.476
IFNAR2	33241945	rs2229207	T	C	MODERATE	0,343	0.079	0.166	0.178	0.084	0.130
IFNAR2	33262573	rs1131668	G	A	MODIFIER	0,795	0.376	0.453	0.583	0.365	0.476
IFNAR2	33262760	rs9984273	C	T	MODIFIER	1	0.594	0.848	0.918	0.687	0.767
IFNAR2	33245424	rs2252639	A	T	MODIFIER	0,285	1	1	1	1	1
IFNAR2	33244908	rs2834158	T	C	MODIFIER	0,333	0.807	0.512	0.416	0.663	0.524
IFNAR2	33230489	rs17860118	G	T	MODIFIER	0,187	0.076	0.118	0.139	0.0864	0.120
IFNAR2	33262748	rs3216172 rs397789038	G	GT	MODIFIER	0,181	0.366	0.467	0.594	0.33	0.484
IFNAR2	33230629	rs12482556	T	C	MODIFIER	0,714	-	-	-	-	-
IFNAR2	33252224	rs56197608	C	T	MODIFIER	0	-	0.003	0.021	0.0002	0.0001
ELF5	34489994	rs2231825	C	A	MODIFIER	0,046	0.116	0.186	0.027	0.429	0.301
ELF5	34505605	rs2231821	A	G	MODIFIER	0,043	0.059	0.155	0.026	0.316	0.265
ELF5	34493816	rs556840829	A	G	MODIFIER	0,05	-	-	-	-	-
ELF5	34511455	rs28395819	A	C	MODIFIER	0,083	-	-	-	-	-
ELF5	34493394	rs737254	T	G	MODIFIER	0,040	0.286	0.327	0.051	0.467	0.339
LIMD1	45595761	rs267237	C	T	MODERATE	0,937	0.584	0.721	0.938	0.562	0.675
LIMD1	45674275	rs2742409	G	A	MODIFIER	0,083	-	-	-	-	-
LIMD1	45674196	rs2578679	G	A	MODIFIER	0,5	-	-	-	-	-
OAS1	112919637	rs2660	G	A	MODERATE	1	0.934	0.840	0.791	0.679	0.707
OAS1	112911065	rs1131454	G	A	MODERATE	0,645	0.228	0.703	0.614	0.575	0.572
OAS1	112931839	rs7968145	C	T	MODIFIER	1	0.989	0.955	1	0.924	0.955
OAS1	112919404	rs1131476	G	A	MODIFIER	1	0.934	0.8318	0.784	0.659	0.698

OAS1	112919388	rs10774671	G	A	MODIFIER	0,833	0.424	0.799	0.783	0.654	0.701
OAS1	112931954	rs7967461	G	C	MODIFIER	1	0.917	0.739	0.76	0.575	0.692
OAS1	112931910	rs11352835	TA	T	MODIFIER	1	0.916	0.738	0.769	0.59	0.691
OAS1	-	rs1051042	G	C	MODIFIER	1	0.934	0.831	0.784	0.657	0.693
UPK1A	35668466	rs2267586	T	G	MODERATE	0,179	0.0715	0.1157	0.2241	0.1007	0.1026
UPK1A	35676195	rs747589460	CTTT	C	MODIFIER	0,033	0.0227	0.037	0.0417	0.0587	0.0811
UPK1A	35673206	rs75289222	G	T	MODIFIER	0,035	0.0653	0.115	0.1029	0.1016	0.0845
UPK1A	35678012	rs2285421	T	C	MODIFIER	0,3	0.8012	0.5853	0.6216	0.5365	0.7368

[a]Reference Allele; [b]Variant Allele; *Variants without described SNP; (-) – No annotation; INDG: Indigenous populations; AFR: African populations; AMR: American populations; EAS: East Asian population; EUR: European population; SAS: South Asian population.

The 45 variants described in Table 1 were characterized by information such as chromosomal position, SNP ID, nucleotide alteration and classification by SNPeff, excluding low-impact ones (except the new variants). A high-impact variant was identified in the ARHGAP27 gene (rs201721078) at position 45404053 with a frequency of 21.8% in the INDG population, being rare in the rest of the world.

Thirty-six impact modifier variants were also identified in the 7 genes studied: three in AQP3 (rs2231235, rs2231231 and rs12555686), eight ARHGAP27 (rs7222444, rs62064597, rs115993362, rs35327136, rs7213200, rs142163608, rs184103721 and rs35389313), five in *ELF5* (rs2231825, rs2231821, rs556840829, rs28395819 and rs737254), eight *IFNAR2* (rs1131668, rs9984273, rs2252639, rs2834158, rs17860118, rs3216172 rs397789038, rs12482556, rs56197608 and a new variant), three in LIMD1 (rs2742409, rs2578679 and a new variant), six in OAS1 (rs7968145, rs1131476, rs10774671, rs7967461, rs11352835 and rs1051042) e three in UPK1A (rs747589460, rs75289222 and rs2285421).

Of the variants found in this study, rs12555686 and rs2231235 (AQP3) and rs56197608 (IFNAR2) did not have an allelic frequency described in the INDG population. However, further studies are needed to confirm such results. In the IFNAR2 gene, rs397789038 and rs12482556 showed 18.1% and 71.4% allele frequency in the INDG population, respectively. In the OAS1 gene, (rs11352835) 100% frequency was found and in UPK1A (rs747589460) 33%. All these variants had a modifying impact.

We also identified 3 new variants in Amazonian indigenous people (Table 2). The first in the AQP3 gene at position 33442882 with low impact and allele frequency of 8.3%.The second variant was identified in the IFNAR2 gene at position 33262799 with a moderate impact and a frequency of 8.3%, and the third was found in the LIMD1 gene at position 45676789 with a modifying impact and frequency of 6.6%.

Table 2 – Description of new variants.

Gene	Position	SNP ID	Ref[a]	Var[b]	Allele Frequence	Impact Predictedby SNPeff
AQP3	33442882	*	G	A	8,3%	LOW
IFNAR2	33262799	*	C	A	8,3%	MODERATE
LIMD1	45676789	*	G	T	6,6%	MODIFIER

[a]Reference Allele; [b]Variant Allele; *Variants without described SNP.

In addition, 10 variants of moderate impact were also found in 5 of the 7 genes studied: 3 in the *ARHGAP27* gene (rs2959953, rs12949256 and rs117139057), 3 in the *IFNAR2* gene (rs1051393, rs2229207 and the presence of a new variant), 1 *LIMD1* (rs267237), 2 in OAS1 (rs1131454 and rs2660) e 1 in *UPK1A* (rs2267586).

Among the variants identified, 17 showed significant differences when compared to other world populations (AFR, AMR, SAS, EAS and EUR), even in East Asians who have greater genetic similarity with indigenous people.: 2 variants were identified in the *ARHGAP27* gene (rs201721078 and rs2959953), 3 present in *AQP3* (rs2228332, rs591810 and rs2231231), 1 in the *UPK1A* gene (rs2285421), 4 in *OAS1* (rs2660, rs7967461, rs11352835 and rs1051042) and 7 found in the *IFNAR2* gene (rs1051393, rs2229207, rs1131668, rs9984273, rs2834158, rs149186597 rs79402470 and rs3216172 rs397789038) (Table 3). The frequencies of the other variants did not show significant differences between the INDG population and the other continental populations (Supplementary Table S2).

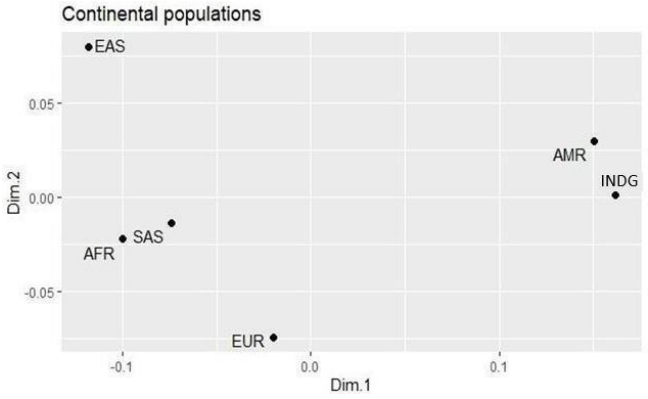
Table 3 - Comparison of the allele frequency of Indigenous and the world population (AFR, AMR, EUR, EAS, and SAS).

Gene	SNP ID	INDG vs. AFR *	INDG vs. AMR *	INDG vs. EAS *	INDG vs. EUR *	INDG vs. SAS *
ARHGAP27	rs201721078	5.129×10^{-15}	1.542×10^{-5}	NA	1.381×10^{-14}	NA
AQP3	rs2228332	0.00018	3.693×10^{-5}	5.348×10^{-5}	4.293×10^{-9}	9.082×10^{-10}
AQP3	rs591810	3.963×10^{-5}	0.00075	1.031×10^{-5}	3.214×10^{-7}	1.087×10^{-8}
AQP3	rs2231231	0.00117	0.00010	5.442×10^{-5}	1.620×10^{-7}	2.264×10^{-9}
UPK1A	rs2285421	3.016×10^{-16}	3.180×10^{-5}	1.341×10^{-6}	0.00030	1.562×10^{-11}
OAS1	rs2660	0.02583	7.618×10^{-5}	1.374×10^{-6}	1.525×10^{-10}	2.482×10^{-9}
OAS1	rs7967461	0.01055	3.650×10^{-8}	1.260×10^{-7}	1.118×10^{-14}	6.043×10^{-10}
OAS1	rs11352835	0.01087	3.432×10^{-8}	2.366×10^{-7}	3.623×10^{-14}	6.043×10^{-10}
OAS1	rs1051042	0.02583	4.458×10^{-5}	7.665×10^{-7}	3.220×10^{-11}	5.950×10^{-10}
ARHGAP27	rs2959953	6.052×10^{-7}	4.116×10^{-5}	0.03356	1.238×10^{-7}	6.31360×10^{-6}
IFNAR2	rs1051393	1.824×10^{-21}	1.053×10^{-5}	0.00258	1.215×10^{-11}	3.942×10^{-6}
IFNAR2	rs2229207	2.456×10^{-8}	0.00186	0.003	1.035×10^{-7}	6.205×10^{-5}
IFNAR2	rs1131668	1.229×10^{-10}	2.771×10^{-7}	0.00098	7.147×10^{-11}	9.955×10^{-7}
IFNAR2	rs9984273	3.001×10^{-14}	0.00013	0.00938	3.134×10^{-10}	2.268×10^{-7}
IFNAR2	rs2834158	5.32×10^{-15}	0.00664	0.22382	3.818×10^{-7}	0.00342
IFNAR2	rs149186597 rs79402470	8.004×10^{-5}	0.00602	1.575×10^{-5}	0.04188	0.04602
IFNAR2	rs3216172 rs397789038	0.00380	2.699×10^{-5}	4.368×10^{-10}	0.02185	3.997×10^{-6}

INDG: Indigenous populations; AFR: African populations; AMR: American populations; EAS: East Asian population; EUR: European population; SAS: South Asian population; *p-value defined by Fisher's exact test. Bold characters indicate a significant difference (p-value* < 0.05).

The genetic differences between the populations (Figure 1) of the study were analyzed using the multidimensional scale graph (MDS), based on the Fisher fixation test of the genetic variants. Multidimensional scale analysis (MDS) identified greater similarity between the INDG and AMR populations, mainly due to the influence of indigenous peoples on the AMR populations. A difference was also identified between EUR, AFR, SAS and the EAS population that showed a greater difference when compared to the other populations.

Figure 1 - Differences in the allelic frequencies of the variants studied in continental populations and the Indigenous population, ploted in MDS.



3. Discussion

The occurrence of new infectious diseases poses a threat to indigenous populations. Since the colonial period, these peoples have suffered from devastating epidemics, such as measles, flu and tuberculosis [10-11]. One of the contributing factors to the occurrence of these diseases may have been influenced by the unique genetic profile of the indigenous population, as well as by genetic mutations not yet known (or rare) [11]. Geographical isolation and the existence of consanguineous marriages are some of the factors that can favor the differentiation of the allele frequency of these populations when compared to other world populations [12-13].

GWAS studies have been developed to identify genes and genetic variants associated with the severity of COVID-19. According to an investigation carried out in Spanish populations, the locus 3p21.31 (*SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6* and *XR1*) and the locus 9q34.21 (*ABO*), showed a strong association with the most severe forms of the disease [7]. These results corroborate those observed in another study carried out in Amazonian Amerindian populations [10] in which the same genes were investigated. However, the identified genetic variants showed no association with more critical COVID-19 conditions in the indigenous population [10].Our study identified the presence of three new variants that may be potentially related to the severity of COVID-19 indigenous peoples of the Brazilian Amazon. The first genetic variant was identified in the AQP3 gene, a gene expressed in a variety of cell types [14-15-16]. Currently, in a GWAS study, 49 variants were associated with the most severe forms of COVID-19, showing that the *AQP3* gene had an intense relationship with the most critical picture of infection, a fact that corroborates other studies [17-9].

The same genes we investigated were also analyzed in a research developed in 34 Spanish hospitals with 11,939 positive cases of COVID-19, suggesting the association of these with the most severe forms of SARS-CoV-2 infection. A difference in the severity of the disease between genders was identified in this study [9]. Due to the greater propensity to develop a more critical picture of SARS-CoV-2 infection in males, the study investigated the relationship of androgenic hormones with the severity of COVID-19 [9]. According to these investigations, these hormones were pointed out as one of the factors responsible for the highest mortality rates of this disease in men [15]. Such data suggest that the genes: *AQP3*, *ARHGAP27, ELF5, IFNAR2, LIMD1, OAS1* and *UPK1A*, related to sexual differences in Spaniards for the development of more severe forms of COVID-19, may also be involved with the androgenic hormone pathways.

We also identified a new variant in the *IFNAR2* gene that is closely related to the severity of COVID-19 [18]. A study developed in the United Kingdom reported an association of this gene with the severity of the disease [19]. In addition, a last new variant has also been identified in the *LIMD1* gene that is involved in several cellular processes [20]. Your cellular location may play a key role in the progression of this type of cancer [21-22].

Finally, we identified a high-impact variant present in the *ARHGAP27* gene (rs201721078) never before associated with COVID-19. The identification of the three new variants and the high-impact one present in the Amerindians in

rs201721078, suggest a relationship with the most serious cases of COVID-19.

This study investigated genes potentially related to the severity of COVID-19 in an Amazonian indigenous. The results found in this study suggest the urgency of more effective research that proves the impact of these new and high-impact variants in patients with SARS-CoV-2 infection in the indigenous populations of the Amazon, aiming to elucidate the biological role of these variants in the severity of COVID -19 in indigenous peoples and contributing to the development of personalized medicine that respects the particularities of the studied population.

4. Materials and methods

4.1. Consent and ethics

This study was approved by the National Research Ethics Committee (CONEP) and the Research Ethics Committee of the Center for Tropical Medicine with the opinion 20654313.6.0000.5172 and CAAE 33934420.0.0000.5634. The representatives of the groups participating in the study were informed about the stages of the study signed on the Free and Informed Consent Form (TCLE). Their materials were collected according to the Declaration of Helsinki.

4.2. Study population

The study was carried out by a blood collection of 64 healthy indigenous people from the Brazilian Amazon region, belonging to the original groups: Asurini do Tocantins, Asurini do Xingu, Araweté, Arara, Juruna, Awa-Guajá, Kayapó/ Xikrin, Munduruku, Karipuna, Phurere, Wajãpi and Zo'é. Blood collection was performed before the pandemic period.

Information from markers indicative of ancestry (AIMs) was obtained to confirm the ancestry and the mixture between continental populations (European, African and Asian) in three multiplex PCR reactions [23-24-25]. Electrophoresis was used to analyze the amplicons in the sequencer ABI Prism 3130 and GeneMapper ID v.3.2. In addition, the proportions of the individuals were analyzed in the STRUCTURE v.2.33 software. The allele frequency data of the INDG population were obtained by the allele count and compared with data from 5 other continental populations (AFR, AMR, EAS, EUR and SAS) found in the Project 1000 Genomes database (http://www.1000genomes.org).

4.3. DNA extraction and preparation of the exome library

The sague collection of each participant in this study was carried out in 5 ml tubes. Subsequently, this material was extracted with the Roche Applied Science DNA extraction kit (Roche, Penzberg, Germany), according to the manufacturer's instructions. The samples were quantified in NanoDrop1000 to verify the integrity of the genetic material. The exoma library was prepared with the help of the Nextera Rapid Capture Exome kit (Illumina, San Diego, CA, USA) and the SureSelect Human All Exon V6 kit (Agilent, Santa Clara, CA, USA). The sequencing step was developed on the NextSeq 500 (Illumina, San Diego, CA, EUA) using the NextSeq 500 v2 300 high production cycle kit (Illumina, San Diego, CA, EUA).

4.4. Bioinformatics analysis

First, the low-quality sequences were eliminated. Subsequently, the sequences that showed good quality were mapped and aligned according to a reference genome (GRCh38) using the BWA v.0.7 software. The variants were identified in GATK v. 3.2 and noted in the Viewer of Variants software (ViVa, Federal University of Rio Grande do Norte, Natal, Brazil). Other databases were also used: npEff v.4.3. T, Ensembl Variant Effect Predictor (Ensembl version 99) and ClinVar (v.2018–10). The SIFT (v.6.2.1), PolyPhen-2 (v.2.2), LRT (November 2009), Mutation Assessor (v.3.0), Mutation Taster (v.2.0), FATHMM (v.2.3), PROVEAN (v.1.1.3), MetaSVM (v.1.0), M-CAP (v.1.4) and FATHMM-MKL. Bioinformatics analyses were performed as described in Ribeiro-Dos-Santos et al. [26] and Rodrigues et al. [25].

4.5. Statistical analyses

In the statistical analysis, two tests were used: the first was Fisher's exact test, to differentiate the frequencies between the populations of the world. The results obtained were considered statistically significant when $p \leq 0.05$. Subsequently, the Wright fixation index (FST) was used to verify the interpopulation variability of genetic variants. The statistical analyses of this study were developed in the software Arlequin v.3.518 and R Studio to develop the graphic data.

4.6. Selection of variants

Three basic criteria were used for the selection of variants: initially, those were selected that had at a) at least 10 coverage readings (fastx_tools v.0.13-http://hannonlab.cshl.edu/fastx_toolkit/; accessed on January 20, 2022); b) an allelic frequency described in the continental populations of the 1000 Genomes Project Consortium [27]; c) variants must have the modifying impact, moderate or high, as classified by SNPeff [28]. Seven genes were used in this study (AQP3, ARHGAP27, ELF5, IFNAR2, LIMD1, OAS1 and UPK1A), selected based on the study by Cruz et al., [9].

5. References

[1] Lu H, Stratton CW, Tang YW. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. J Med Virol. 92(4):401-402. (2020). doi: 10.1002/jmv.25678. [PubMed]

[2] 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, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet.15;395(10223):497-506. (2020). doi: 10.1016/S0140-6736(20)30183-5. [PubMed].

[3] Dhama, K., Patel, S. K., Pathak, M., Yatoo, M. I., Tiwari, R., Malik, Y. S., et al. An update on SARS-CoV-2/COVID-19 with particular reference to its clinical pathology, pathogenesis, immunopathology and mitigation strategies. Travel. Med. Infect. Dis. 37:101755. (2020). doi: 10.1016/j.tmaid.2020.101755

[4] Cucinotta, D., and Vanelli, M. WHO declares COVID-19 a pandemic. Acta Biomed. 91, 157–160. (2020). doi: 10.23750/abm.v91i1.9397

[5] World Health Organization. WHO Coronavirus (COVID-19) Dashboard. Available online: https://covid19.who.int (accessed on 30 April 2023)

[6] da Silva Torres, M. K., Bichara, C. D. A., de Almeida, M. N. D. S., Vallinoto, M. C., Queiroz, M. A. F., Vallinoto, I. M. V. C., Dos Santos, E. J. M., de Carvalho, C. A. M., & Vallinoto, A. C. R. The Complexity of SARS-CoV-2 Infection and the COVID-19 Pandemic. Frontiers in microbiology, (2022) 13, 789882. https://doi.org/10.3389/fmicb.2022.789882

[7] Ellinghaus, D. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. N. Engl. J. Med. 2020, 383, 1522–1534

[8] Shelton, J.F.; Shastri, A.J.; Ye, C.; Weldon, C.H.; Filshtein-Sonmez, T.; Coker, D.; Symons, A.; Esparza-Gordillo, J.; Aslibekyan, S.; Auton, A.; et al. Trans-ancestry analysis reveals genetic and nongenetic associations with COVID-19 susceptibility and severity. Nat. Genet. 2021, 53, 801–808

[9] Cruz R, Almeida SD, Heredia ML, Quintela I, Ceballos FC, Pita G, Lorenzo-Salazar JM, González-Montelongo R, Gago-Domínguez M, Porras MS, Castaño JAT, Nevado J, Aguado JM, Aguilar C, Aguilera-Albesa S, Almadana V, Almoguera B, Alvarez N, Andreu-Bernabeu Á, Arana-Arri E, Arango C, Arranz MJ, Artiga MJ, Baptista-Rosas RC, Barreda-Sánchez M, Belhassen-Garcia M, Bezerra JF, Bezerra MAC, Boix-Palop L, Brion M, Brugada R, Bustos M, Calderón EJ, Carbonell C, Castano L, Castelao JE, Conde-Vicente R, Cordero-Lorenzana ML, Cortes-Sanchez JL, Corton M, Darnaude MT, De Martino-Rodríguez A, Campo-Pérez V, Bustamante AD, Domínguez-Garrido E,

Luchessi AD, Eirós R, Sanabria GME, Fariñas MC, Fernández-Robelo U, Fernández- Rodríguez A, Fernández-Villa T, Gil-Fournier B, Gómez-Arrue J, Álvarez BG, Quirós FGB, González-Peñas J, Gutiérrez-Bautista JF, Herrero MJ, Herrero-Gonzalez A, Jimenez-Sousa MA, Lattig MC, Borja AL, Lopez-Rodriguez R, Mancebo E, Martín- López C, Martín V, Martinez-Nieto O, Martinez-Lopez I, Martinez-Resendez MF, Martinez-Perez Á, Mazzeu JA, Macías EM, Minguez P, Cuerda VM, Silbiger VN, Oliveira SF, Ortega-Paino E, Parellada M, Paz-Artal E, Santos NPC, Pérez-Matute P, Perez P, Pérez-Tomás ME, Perucho T, Pinsach-Abuin ML, Pompa-Mera EN, Porras- Hurtado GL, Pujol A, León SR, Resino S, Fernandes MR, Rodríguez-Ruiz E, Rodriguez-Artalejo F, Rodriguez-Garcia JA, Ruiz-Cabello F, Ruiz-Hornillos J, Ryan P, Soria JM, Souto JC, Tamayo E, Tamayo-Velasco A, Taracido-Fernandez JC, Teper A, Torres-Tobar L, Urioste M, Valencia-Ramos J, Yáñez Z, Zarate R, Nakanishi T, Pigazzini S, Degenhardt F, Butler-Laporte G, Maya-Miles D, Bujanda L, Bouysran Y, Palom A, Ellinghaus D, Martínez-Bueno M, Rolker S, Amitrano S, Roade L, Fava F, Spinner CD, Prati D, Bernardo D, Garcia F, Darcis G, Fernández-Cadenas I, Holter JC, Banales JM, Frithiof R, Duga S, Asselta R, Pereira AC, Romero-Gómez M, Nafría- Jiménez B, Hov JR, Migeotte I, Renieri A, Planas AM, Ludwig KU, Buti M, Rahmouni S, Alarcón-Riquelme ME, Schulte EC, Franke A, Karlsen TH, Valenti L, Zeberg H, Richards B, Ganna A, Boada M, Rojas I, Ruiz A, Sánchez P, Real LM; SCOURGE Cohort Group; HOSTAGE Cohort Group; GRA@CE Cohort Group, Guillen-Navarro E, Ayuso C, González-Neira A, Riancho JA, Rojas-Martinez A, Flores C, Lapunzina P, Carracedo Á. Novel genes and sex differences in COVID-19 severity. Hum Mol Genet. (2022). 16:ddac132. doi: 10.1093/hmg/ddac132. [PubMed]

[10] Pastana, L. F., Silva, T. A., Gellen, L. P. A., Vieira, G. M., de Assunção, L. A., Leitão, L. P. C., da Silva, N. M., Coelho, R. C. C., de Alcântara, A. L., Vinagre, L. W. M. S., Rodrigues, J. C. G., Borges Leal, D. F. D. V., Fernandes, M. R., de Souza, S. J., Kroll, J. E., Ribeiro-Dos-Santos, A. M., Burbano, R. M. R., Guerreiro, J. F., de Assumpção, P. P., Ribeiro-Dos-Santos, Â. C., … Dos Santos, N. P. C. The Genomic Profile Associated with Risk of Severe Forms of COVID-19 in Amazonian. Native American Populations. Journal of personalized medicine, (2022) 12(4), 554. https://doi.org/10.3390/jpm12040554

[11] Walker RS, Sattenspiel L, Hill KR. Mortality from contact-related epidemics among indigenous populations in Greater Amazonia. Sci Rep. 2015 Sep 10;5:14032. doi: 10.1038/srep14032

[12] Salzano, F.M. Fatores dete rminísticos e estocásticos no processo microevolucionário humano. Actas V Congr. Latinoam. Genet. 1982, 1, 81–89

[13] De Brito Vargas, L.; Beltrame, M.H.; Ho, B.; Marin, W.M.; Dandekar, R.; Montero-Martín, G.; Fernández-Viña, M.A.; Hurtado, A.M.; Hill, K.R.; Tsuneto, L.T.; et al. Remarkably Low KIR and HLA Diversity in Amerindians Reveals Signatures of Strong Purifying Selection Shaping the Centromeric KIR Region. Mol. Biol. Evol. 2022, 39, msab298. [CrossRef]

[14] Michałek, K., 2016. Aquaglyceroporins in the kidney: present state of knowledge and prospects. J. Physiol. Pharmacol. 67, 185–193

[15] Li, C., Wang, W., 2017. Molecular biology of aquaporins. Adv. Exp. Med. Biol. 969, 1–34

[16] Oberska P, Jedrzejczak-Silicka M, Michałek K, Grabowska M. Initial assessment of suitability of MCF-7 and HepG2 cancer cell lines for AQP3 research in cancer biology. Acta Histochem. 2021 May;123(4):151716. doi10.1016/j.acthis.2021.151716. Epub 2021 Apr 29. PMID: 33933702

[17] Pairo-Castineira E, Rawlik K, Bretherick AD, Qi T, Wu Y, Nassiri I, McConkey GA, Zechner M, Klaric L, Griffiths F, Oosthuyzen W, Kousathanas A, Richmond A, Millar J, Russell CD, Malinauskas T, Thwaites R, Morrice K, Keating S, Maslove D, Nichol A, Semple MG, Knight J, Shankar-Hari M, Summers C, Hinds C, Horby P, Ling L, McAuley D, Montgomery H, Openshaw PJM, Begg C, Walsh T, Tenesa A, Flores C, Riancho JA, Rojas-Martinez A, Lapunzina P; GenOMICC Investigators; SCOURGE Consortium; ISARICC Investigators; 23andMe COVID-19

Team; Yang J, Ponting CP, Wilson JF, Vitart V, Abedalthagafi M, Luchessi AD, Parra EJ, Cruz R, Carracedo A, Fawkes A, Murphy L, Rowan K, Pereira AC, Law A, Fairfax B, Hendry SC, Baillie JK. GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19. Nature. 2023 May;617(7962):764-768. doi: 10.1038/s41586-023-06034-3. Epub 2023 May 17. PMID: 37198478; PMCID: PMC10208981

[18] Fricke-Galindo I, Martínez-Morales A, Chávez-Galán L, Ocaña-Guzmán R, Buendía-Roldán I, Pérez-Rubio G, Hernández-Zenteno RJ, Verónica-Aguilar A, Alarcón-Dionet A, Aguilar-Duran H, Gutiérrez-Pérez IA, Zaragoza-García O, Alanis-Ponce J, Camarena A, Bautista-Becerril B, Nava-Quiroz KJ, Mejía M, Guzmán- Guzmán IP, Falfán-Valencia R. IFNAR2 relevance in the clinical outcome of individuals with severe COVID-19. Front Immunol. 2022 Jul 29;13:949413. doi: 10.3389/fimmu.2022.949413. PMID: 35967349; PMCID: PMC9374460

[19] Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al. Genetic mechanisms of critical illness in covid-19. Nature (2020) 591:92–8. doi: 10.1038/ s41586-020-03065-y

[20] Zhou J, Zhang L, Zhou W, Chen Y, Cheng Y, Dong J. LIMD1 phosphorylation in mitosis is required for mitotic progression and its tumor-suppressing activity. FEBS J. 2019 Mar;286(5):963-974. doi: 10.1111/febs.14743. Epub 2019 Jan 16. PMID: 30600590

[21] Sharp TV, Munoz F, Bourboulia D, Presneau N, Darai E, Wang HW, Cannon M, Butcher DN, Nicholson AG, Klein G, Imreh S, Boshoff C LIM domains-containing protein 1 (LIMD1), a tumor suppressor encoded at chromosome 3p21.3, binds pRB and represses E2F-driven transcription. Proc Natl Acad Sci USA 2004;101:16531–6

[22] Spendlove I, Al-Attar A, Watherstone O, Webb TM, Ellis IO, Longmore GD, Sharp TV. Differential subcellular localisation of the tumour suppressor protein LIMD1 in breast cancer correlates with patient survival. Int J Cancer. 2008 Nov15;123(10):2247-53. doi: 10.1002/ijc.23851. PMID: 18712738

[23] Leal, D.F.D.V.B.; Da Silva, M.N.S.; Fernandes, D.C.R.D.O.; Rodrigues, J.C.G.; Barros, M.C.D.C.; Pinto, P.D.D.C.; Pastana, L.F.; Da Silva, C.A.; Fernandes, M.R.; De Assumpção, P.P.; et al. Amerindian genetic ancestry as a risk factor for tuberculosis in an amazonian population. PLoS ONE 2020, 15, e0236033. [CrossRef] [PubMed]

[24] De Ramos, B.R.A.; D'Elia, M.P.B.; Amador, M.A.T.; Santos, N.P.C.; Santos, S.E.B.; da Castelli, E.C.; Witkin, S.S.; Miot, H.A.; Miot, L.D.B.; da Silva, M.G. Neither self-reported ethnicity nor declared family origin are reliable indicators of genomic ancestry. Genetica 2016, 144, 259–265. [CrossRef]

[25] Rodrigues, J.C.G.; de Souza, T.P.; Pastana, L.F.; Ribeiro dos Santos, A.M.; Fernandes, M.R.; Pinto, P.; Wanderley, A.V.; De Souza, S.J.; Kroll, J.E.; Pereira, A.L.; et al. Identification of NUDT15 gene variants in Amazonian Amerindians and admixed individuals from northern Brazil. PLoS ONE 2020, 15, e0231651 [CrossRef]

[26] Ribeiro-dos-Santos, A.M.; Vidal, A.F.; Vinasco-Sandoval, T.; Guerreiro, J.; Santos, S.; Ribeiro-dos-Santos, Â. Exome Sequencing of Native Populations from the Amazon Reveals Patterns on the Peopling of South America. Front. Genet. 2020, 11, 548507. [CrossRef] [PubMed]

[27] Pinto, P.; Salgado, C.; Santos, N.P.C.; Santos, S.; Ribeiro-dos-Santos, Â. Influence of Genetic Ancestry on INDEL Markers of NFKβ1, CASP8, PAR1, IL4 and CYP19A1 Genes in Leprosy Patients. PLoS Negl. Trop. Dis. 2015, 9, e0004050. [CrossRef]

[28] Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 2012 6, 80–92. [CrossRef] [PubMed]

Acknowledgments: We thank the Federal University of Pará (UFPA); the Center for Research in Oncology (NPO/UFPA), the Graduate Program in Genetics and Molecular Biology (PPGM/UFPA) and the Laboratory of Human and Medical Genetics (LGHM/UFPA). We also thank the contributions and ethical evaluation of indigenous rights of Josicleyce Pereira.

Author Contributions: Data curation: R.d.C.C.C., idealization, contextualization, investigation and preparation of the original writing; C.L.L.P.M., conceptualization and preparation of the original writing; L.F.P., statistical analysis; J.C.G.R., statistical analysis, methodology and software; K.E.C., original writing preparation; A.d.N.C.-P., writing-review and editing; L.P.A.G., translation and preparation of the original script; F.C.A.M., adjustments to the rules and preparation of the original script; M.C.L.C., original writing preparation; L.A.d.A., visualization; N.M., original writing preparation; E.E.B.P.: original writing preparation; A.M.R.d.S., methodology; A.C.R.d.S. methodology and research; R.M.R.B., financial acquisition, writing proofreading and editing; S.J.d.S., methodology and software; J.F.G., data curation and original writing preparation; P.P.d.A., proofreading and editing of writing; S.E.B.d.S., conceptualization, financial acquisition, proofreading, and writing editing; M.F.R., preparation of original writing, revision and editing of writing; N.P.C.d.S., conceptualization, revision and editing of writing.

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

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki and approved by the National Committee for Ethics in Research (CONEP), and the Research Ethics Committee of the Center for Tropical Medicine with the opinion 20654313.6.0000.5172 and CAAE 33934420.0.0000.5634.

Data Availability Statement: The data obtained from the public domain are available at gnomAD (broadinstitute.org), and the sequencing data of the Amazonian Indigenous populations are available at the ENA database under the accession number PRJEB35045.

Supplementary Information: Supplementary Table 1 online provides the 66 variants distributed in the AQP3, ARHGAP27, ELF5, IFNAR2, LIMD1, OAS1 and UPK1A genes. Supplementary Table 2 online are present the other variants that did not present a significant difference between the INDG population and the other continental populations.

4. CONCLUSÕES INTEGRADORAS

O desenvolvimento de estudos genômicos torna possível a detecção de variantes associadas com a gravidade da COVID-19, modificando o cenário precário de pesquisas capazes de analisar o perfil molecular de pacientes mais complexos, com particularidades pouco conhecidas, o que pode promover um reflexo positivo sobre o diagnóstico e tratamento dessa doença dentro dessas populações.

No entanto, ainda há um grande déficit de estudos genômicos dentro dessas populações em geral. Estudos que busquem aumentar as descobertas genéticas e relevâncias clínicas refletindo positivamente na saúde pública da população indígena. Frente a isso, em Pastana *et al.* (2022), nosso grupo de pesquisa analisou variantes genéticas possivelmente associadas com as formas mais graves da COVID-19. Porém, não foram encontradas variantes relacionadas com aseveridade dessa doença em populações indígenas da Amazônia brasileira. Nossos resultados identificaram 15 SNPs com impacto moderado presente nos genes *ABO,CXCR6, FYCO1* e *SC6A20.* Adicionalmente, também identificamos 18SNPs com diferença na frequência alélica e três novas variantes identificadas nos genes *ABO,FYCO1* e *LZTFL1.*

Em nosso estudo, foram encontradas 17 variantes com diferença significativa na população indígena quando comparada com as demais populações estudadas. Destas, uma de alto impacto foi encontrada no gene *ARHGAP27* (rs201721078). Três novas variantes também foram identificadas presentes nas posições: 33442882 (*AQP3*), 33262799 (*IFNAR2*) e 45676789 (*LIMD1*). Através desses dados, também pode ser sugerido na análise em escala multidimensional (MDS) que as populações indígenas apresentam similaridade entre as populações indígenas e as de origem Latina, provavelmente em virtude da influência dos povos indígenas nas populações latinas. Também identificamos diferenças entre essas populações com as de origem Europeia, Africana, Sul Asiático e Leste Asiático.

Os resultados encontrados nesse estudo reforçam a ideia do perfil genético único da população indígena e ressaltam a importância de maiores investigações que consigam elucidar o comportamento da COVID-19 nesses grupos populacionais. Estudos com maior número populacional que compreendam o real impacto das variantes aqui identificadas são fundamentais para confiabilizar nossos resultados.

Assim, tais dados podem contribuir para o cenário clínico dessa doença em populações indígenas. No que concerne isso, este presente estudo contribui para a elucidação da importância da genômica na compreensão da associação de variantes genéticas com as formas mais severas da COVID-19 em populações indígenas da Amazônia brasileira.

5. REFERÊNCIAS

Alvim, R. D. O. et al. Body mass index, waist circumference, body adiposity index, and risk for type 2 diabetes in two populations in Brazil: general and Amerindian. PloS one, v. 9, n. 6, 2014.

Araujo DB, Machado RRG, Amgarten DE, Malta FM, de Araujo GG, Monteiro CO, . . . Durigon EL. SARS-CoV-2 isolation from the first reported patients in Brazil and establishment of a coordinated task network. Mem Inst Oswaldo Cruz. 23;115:e200342, out. 2020.

Arnaiz-Villena, A. et al. Amerindians show association to obesity with adiponectin gene SNP45 and SNP276: population genetics of a food intake control and “thrifty” gene. Molecular biology reports, v. 40, n. 2, p. 1819–1826, 2013.

Baloch S, Baloch MA, Zheng T, Pei X. The Coronavirus Disease 2019 (COVID-19) Pandemic. Tohoku J Exp Med. 250(4):271-278, abr. 2020.

Barbosa, C. C. et al. Anthropometric and metabolic profile of a Brazilian Amerindian group: The Xikrin (Mebengôkre). American Journal of Human Biology, v. 31, n. 4, 2019.

Chan J. F. W., Kai-Wang K., Tse H., Jin D. Y., Yuen K. Y. (2013a). Interspecies transmission and emergence of novel viruses: lessons from bats and birds. *Trend Microbiol.* 21 544–555. 10.1016/j.tim.2013.05.005 [PMC free article] [PubMed] [CrossRef] [Google Scholar]

Chan J. F. W., Lau S. K. P., To K. K. W., Cheng V. C. C., Woo P. C. Y., Yuen K. Y., et al. (2015b). Middle east respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease. *Clin. Microbiol. Rev.* 2 465–522. 10.1128/CMR.00102-14 [PMC free article] [PubMed] [CrossRef] [Google Scholar]

Chan JWM, Ng CK, Chan YH, et al. Short term outcome and risk factors for adverse clinical outcomes in adults with severe acute respiratory syndrome (SARS). Thorax 2003c; 58: 686–89.

Cruz R, Diz-de Almeida S, López de Heredia M, Quintela I, Ceballos FC, Pita G, Lorenzo-Salazar JM, González-Montelongo R, Gago-Domínguez M, Sevilla Porras M, Tenorio Castaño JA, Nevado J, Aguado JM, Aguilar C, Aguilera-Albesa S, Almadana V, Almoguera B, Alvarez N, Andreu-Bernabeu Á, Arana-Arri E, Arango C, Arranz MJ, Artiga MJ, Baptista-Rosas RC, Barreda-Sánchez M, Belhassen- Garcia M, Bezerra JF, Bezerra MAC, Boix-Palop L, Brion M, Brugada R, Bustos M,Calderón

EJ, Carbonell C, Castano L, Castelao JE, Conde-Vicente R, Cordero- Lorenzana ML, Cortes- Sanchez JL, Corton M, Darnaude MT, De Martino- Rodríguez A, Del Campo-Pérez V, Diaz de Bustamante A, Domínguez-Garrido E, Luchessi AD, Eiros R, Estigarribia Sanabria GM, Carmen Fariñas M, Fernández- Robelo U, Fernández-Rodríguez A, Fernández-Villa T, Gil-Fournier B, Gómez-ArrueJ, González Álvarez B, Gonzalez Bernaldo de Quirós F, González-Peñas J, Gutiérrez-Bautista JF, Herrero MJ, Herrero- Gonzalez A, Jimenez-Sousa MA, Lattig MC, Liger Borja A, Lopez-Rodriguez R, Mancebo E, Martín-López C, Martín V, Martinez- Nieto O, Martinez-Lopez I, Martinez- Resendez MF, Martinez-Perez A, Mazzeu JF, Merayo Macías E, Minguez P, Moreno Cuerda V, Silbiger VN, Oliveira SF, Ortega-Paino E, Parellada M, Paz-Artal E, Santos NPC, Pérez-Matute P, PerezP, Pérez-Tomás ME, Perucho T, Pinsach-Abuin ML, Pompa-Mera EN, Porras- Hurtado GL, Pujol A, Ramiro León S, Resino S, Fernandes MR, Rodríguez-Ruiz E,Rodriguez-Artalejo F, Rodriguez-Garcia JA, Ruiz Cabello F, Ruiz-Hornillos J, RyanP, Soria JM, Souto JC, Tamayo E, Tamayo-Velasco A, Taracido-Fernandez JC, Teper A, Torres-Tobar L, Urioste M, Valencia-Ramos J, Yáñez Z, Zarate R, Nakanishi T, Pigazzini S, Degenhardt F, Butler- Laporte G, Maya- Miles D, BujandaL, Bouysran Y, Palom A, Ellinghaus D, Martínez- Bueno M, Rolker S, Amitrano S, Roade L, Fava F, Spinner CD, Prati D, Bernardo D, Garcia F, Darcis G, Fernández-Cadenas I, Holter JC, Banales JM, Frithiof R, Duga S, Asselta R, Pereira AC, Romero-Gómez M, Nafría-Jiménez B, Hov JR, Migeotte I, Renieri A, Planas AM, Ludwig KU, Buti M, Rahmouni S, Alarcón-Riquelme ME, Schulte EC, Franke A, Karlsen TH, Valenti L, Zeberg H, Richards B, Ganna A, Boada M, de Rojas I, Ruiz A, Sánchez-Juan P, Real LM; SCOURGE Cohort Group; HOSTAGE Cohort Group;GRA@CE Cohort Group; Guillen-Navarro E, Ayuso C, González-Neira A, Riancho JA, Rojas-Martinez A, Flores C, Lapunzina P, Carracedo A. Novel genes and sex differences in COVID-19 severity. Hum Mol Genet. 2022 Nov 10;31(22):3789-3806. doi: 10.1093/hmg/ddac132. PMID: 35708486; PMCID: PMC9652109.

Cucinotta, D., and Vanelli, M. (2020). WHO declares COVID-19 a pandemic. Acta Biomed. 91, 157–160. doi: 10.23750/abm.v91i1.9397

Driggin E, Madhavan MV, Bikdeli B, Chuich T, Laracy J, Biondi-Zoccai G, Brown TS, Der Nigoghossian C, Zidar DA, Haythe J, Brodie D, Beckman JA, Kirtane AJ, Stone GW, Krumholz HM, Parikh SA. Cardiovascular Considerations for Patients, Health Care Workers, and Health Systems During the COVID-19 Pandemic. J Am Coll Cardiol. 2020 May 12;75(18):2352-2371. doi: 10.1016/j.jacc.2020.03.031. Epub

2020 Mar 19. PMID: 32201335; PMCID: PMC7198856.

Drosten C., Günther S., Preiser W., Werf S., Brodt H. R., Becker S., et al. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *Engl. J. Med.* 348 1967–1976. 10.1056/NEJMoa030747 [PubMed] [CrossRef] [Google Scholar]

Dhama K, Patel SK, Pathak M, Yatoo MI, Tiwari R, Malik YS, Singh R, Sah R, Rabaan AA, Bonilla-Aldana DK, Rodriguez-Morales AJ. An update on SARS-CoV-2/COVID-19 with particular reference to its clinical pathology, pathogenesis, immunopathology and mitigation strategies. Travel Med Infect Dis. 2020 Sep-Oct;37:101755. doi: 10.1016/j.tmaid.2020.101755. Epub 2020 May 30. PMID: 32479816; PMCID: PMC7260597.

Severe Covid-19 GWAS Group; Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, Fernández J, Prati D, Baselli G, Asselta R, Grimsrud MM, Milani C, Aziz F, Kässens J, May S, Wendorff M, Wienbrandt L, Uellendahl-Werth F, Zheng T, Yi X, de Pablo R, Chercoles AG, Palom A, Garcia-Fernandez AE, Rodriguez-Frias F, Zanella A, Bandera A, Protti A, Aghemo A, Lleo A, Biondi A, Caballero-Garralda A, Gori A, Tanck A, Carreras Nolla A, Latiano A, Fracanzani AL, Peschuck A, Julià A, Pesenti A, Voza A, Jiménez D, Mateos B, Nafria Jimenez B, Quereda C, Paccapelo C, Gassner C, Angelini C, Cea C, Solier A, Pestaña D, Muñiz-Diaz E, Sandoval E, Paraboschi EM, Navas E, García Sánchez F, Ceriotti F, Martinelli-Boneschi F, Peyvandi F, Blasi F, Téllez L, Blanco-Grau A, Hemmrich-Stanisak G, Grasselli G, Costantino G, Cardamone G, Foti G, Aneli S, Kurihara H, ElAbd H, My I, Galván-Femenia I, Martín J, Erdmann J, Ferrusquía-Acosta J, Garcia-Etxebarria K, Izquierdo-Sanchez L, Bettini LR, Sumoy L, Terranova L, Moreira L, Santoro L, Scudeller L, Mesonero F, Roade L, Rühlemann MC, Schaefer M, Carrabba M, Riveiro-Barciela M, Figuera Basso ME, Valsecchi MG, Hernandez-Tejero M, Acosta-Herrera M, D'Angiò M, Baldini M, Cazzaniga M, Schulzky M, Cecconi M, Wittig M, Ciccarelli M, Rodríguez-Gandía M, Bocciolone M, Miozzo M, Montano N, Braun N, Sacchi N, Martínez N, Özer O, Palmieri O, Faverio P, Preatoni P, Bonfanti P, Omodei P, Tentorio P, Castro P, Rodrigues PM, Blandino Ortiz A, de Cid R, Ferrer R, Gualtierotti R, Nieto R, Goerg S, Badalamenti S, Marsal S, Matullo G, Pelusi S, Juzenas S, Aliberti S, Monzani V, Moreno V, Wesse T, Lenz TL, Pumarola T, Rimoldi V, Bosari S, Albrecht W, Peter W, Romero-Gómez M, D'Amato M, Duga S, Banales JM, Hov JR, Folseraas T, Valenti L, Franke A, Karlsen TH.

Genomewide Association Study of Severe Covid-19 with Respiratory Failure. N Engl J Med. 2020 Oct 15;383(16):1522-1534. doi: 10.1056/NEJMoa2020283. Epub 2020 Jun 17. PMID: 32558485; PMCID: PMC7315890.

Fehr, A. R., and Perlman, S. (2015). Coronaviruses: an overview of their replication and pathogenesis. Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. Methods Mol. Biol. 1282, 1–23. doi: 10.1007/978- 1-4939-2438-7_1

Fricke-Galindo I, Falfán-Valencia R. Genetics Insight for COVID-19 Susceptibility and Severity: A Review. Front Immunol. 2021 Apr 1;12:622176. doi: 10.3389/fimmu.2021.622176. PMID: 33868239; PMCID: PMC8047200.

Güner R, Hasanoğlu I, Aktaş F. COVID-19: Prevention and control measures in community. Turk J Med Sci. 2020 Apr 21;50(SI-1):571-577. doi: 10.3906/sag-2004-146. PMID: 32293835; PMCID: PMC7195988.

Henry BM. COVID-19, ECMO, and lymphopenia: a word of caution. Lancet Respir Med. 2020 Apr;8(4):e24. doi: 10.1016/S2213-2600(20)30119-3. Epub 2020 Mar 13. PMID: 32178774; PMCID: PMC7118650.

Hoffmann, M., Kleine-Weber, H., Krüger, N., Müller, M., Drosten, C., and Pöhlmann, S. (2020). The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells. Cell 181, 271.e8–280.e8. doi: 10.1016/j.cell.2020.02.052

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, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet.15;395(10223):497-506. doi: 10.1016/S0140-6736(20)30183-5. [PubMed]

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE (2010). Os indígenas no Censo Demográfico. Disponível em: http://www.ibge.gov.br/indigenas/indigena_censo2010.pdf. Acessed on July 30, 2023.

International Committee on Taxonomy of Viruses [ICTV] (2019). Available online at: https://talk.ictvonline.org/taxonomy/p/taxonomy-history?taxnode_id=202001847 (Acess 03 de february 2023) [Google Scholar].

International Committee on Taxonomy of Viruses [ICTV] (2020). Available online

at: https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/222/coronaviridae (Acess 03 de february 2023) [Google Scholar]

Jaimes, J., Millet, J., and Whittaker, G. (2020). Proteolytic cleavage of the SARS-CoV-2 spike protein and the role of the novel S1/S2. iScience 23:101212. doi: 10.1016/j.isci.2020.101212

Kim, D., Lee, J., Yang, J., Kim, J., Kim, V., & Chang, H. The Architecture of SARS-CoV-2 Transcriptome. Cell. 181(4), 914-921.e10, 2020.

Lau L. T., Fung Y. W. W., Wong F. P. F., Lin S. S. W., Wang C. R., Li H. L., et al. (2003). A real-time PCR for SARS-coronavirus incorporating target gene pre-amplification. *Biochem. Biophys. Res. Commun.* 312 1290–1296. 10.1016/j.bbrc.2003.11.064 [PMC free article] [PubMed] [CrossRef] [Google Scholar]

Lippi G, Lavie CJ, Sanchis-Gomar F. Cardiac troponin I in patients with coronavirus disease 2019 (COVID-19): Evidence from a meta-analysis. Prog Cardiovasc Dis. 2020 May-Jun;63(3):390-391. doi: 10.1016/j.pcad.2020.03.001. Epub 2020 Mar 10. PMID: 32169400; PMCID: PMC7127395.

Luo M, Gong F, Wang J, Gong Z. The priority for prevention and control of infectious diseases: Reform of the Centers for Disease Prevention and Control - Occasioned by "the WHO chief declares end to COVID-19 as a global health emergency". Biosci Trends. 2023 Jul 11;17(3):239-244. doi: 10.5582/bst.2023.01124. Epub 2023 Jun 22. PMID: 37344393.

Lu H, Stratton CW, Tang YW. (2020). Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. J Med Virol. 92(4):401-402. doi: 10.1002/jmv.25678. [PubMed]

Park J-H, Jang W, Kim S-W, Lee J, Lim Y-S, Cho C-G, et al. The Clinical Manifestations and Chest Computed Tomography Findings of Coronavirus Disease 2019 (COVID-19) Patients in China: A Proportion Meta-Analysis. Clin Exp Otorhinolaryngol (2020) 13:95–105. doi: 10.21053/ceo.2020.00570

Pastana LF, Silva TA, Gellen LPA, Vieira GM, de Assunção LA, Leitão LPC, da Silva NM, Coelho RCC, de Alcântara AL, Vinagre LWMS, Rodrigues JCG, Borges Leal DFDV, Fernandes MR, de Souza SJ, Kroll JE, Ribeiro-Dos-Santos AM, Burbano RMR, Guerreiro JF, de Assumpção PP, Ribeiro-Dos-Santos ÂC, Dos Santos SEB,

Dos Santos NPC. The Genomic Profile Associated with Risk of Severe Forms of COVID- 19 in Amazonian Native American Populations. J Pers Med. 2022 Apr 1;12(4):554. doi: 10.3390/jpm12040554. PMID: 35455670; PMCID: PMC9027999.

Rabaan AA, Al-Ahmed SH, Haque S, Sah R, Tiwari R, Malik YS, . . . Rodriguez-Morales AJ. SARS-CoV-2, SARS-CoV, and MERS-COV: A comparative overview. Infez Med. Ahead Of Print;28(2):174-184, jun. 2020.

Rodriguez-Frias F, Zanella A, Bandera A, Protti A, Aghemo A, Lleo A, Biondi A, Caballero-Garralda A, Gori A, Tanck A, Carreras Nolla A, Latiano A, Fracanzani AL, Peschuck A, Julià A, Pesenti A, Voza A, Jiménez D, Mateos B, Nafria Jimenez B, Quereda C, Paccapelo C, Gassner C, Angelini C, Cea C, Solier A, Pestaña D, Muñiz-Diaz E, Sandoval E, Paraboschi EM, Navas E, García Sánchez F, Ceriotti F, Martinelli-Boneschi F, Peyvandi F, Blasi F, Téllez L, Blanco-Grau A, Hemmrich-Stanisak G, Grasselli G, Costantino G, Cardamone G, Foti G, Aneli S, Kurihara H, ElAbd H, My I, Galván- Femenia I, Martín J, Erdmann J, Ferrusquía-Acosta J, Garcia-Etxebarria K, Izquierdo- Sanchez L, Bettini LR, Sumoy L, Terranova L, Moreira L, Santoro L, Scudeller L, Mesonero F, Roade L, Rühlemann MC, Schaefer M, Carrabba M, Riveiro-Barciela M, Figuera Basso ME, Valsecchi MG, Hernandez-Tejero M, Acosta-Herrera M, D'Angiò M, Baldini M, Cazzaniga M, Schulzky M, Cecconi M, Wittig M, Ciccarelli M, Rodríguez- Gandía M, Bocciolone M, Miozzo M, Montano N, Braun N, Sacchi N, Martínez N, Özer O, Palmieri O, Faverio P, Preatoni P, Bonfanti P, Omodei P, Tentorio P, Castro P, Rodrigues PM, Blandino Ortiz A, de Cid R, Ferrer R, Gualtierotti R, Nieto R, Goerg S, Badalamenti S, Marsal S, Matullo G, Pelusi S, Juzenas S, Aliberti S, Monzani V, Moreno V, Wesse T, Lenz TL, Pumarola T, Rimoldi V, Bosari S, Albrecht W, Peter W, Romero-Gómez M, D'Amato M, Duga S, Banales JM, Hov JR, Folseraas T, Valenti L, Franke A, Karlsen TH. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. N Engl J Med. 2020 Oct 15;383(16):1522-1534. doi: 10.1056/NEJMoa2020283. Epub 2020 Jun 17. PMID: 32558485; PMCID: PMC7315890.

da Silva Torres MK, Bichara CDA, de Almeida MNDS, Vallinoto MC, Queiroz MAF, Vallinoto IMVC, Dos Santos EJM, de Carvalho CAM, Vallinoto ACR. The Complexity of SARS-CoV-2 Infection and the COVID-19 Pandemic. Front Microbiol. 2022. Feb 10;13:789882. doi: 10.3389/fmicb.2022.789882. PMID: 35222327; PMCID: PMC8870622.

Terpos E, Ntanasis-Stathopoulos I, Elalamy I, Kastritis E, Sergentanis TN, Politou M, Psaltopoulou T, Gerotziafas G, Dimopoulos MA. Hematological findings and complications of COVID-19. Am J Hematol. 2020 Jul;95(7):834-847. doi: 10.1002/ajh.25829. Epub 2020 May 23. PMID: 32282949; PMCID: PMC7262337.

Tortorici, M., & Veesler, D. Structural insights into coronavirus entry. Advances In Virus Research. 93-116, 2019.

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

Wang, Q., Zhang, Y., Wu, L., Niu, S., Song, C., Zhang, Z., et al. (2020). Structural and functional basis of SARS-CoV-2 entry by using human ACE2. Cell 181, 894–904.e9. doi: 10.1016/j.cell.2020.03.045

WORLD HEALTH ORGANIZATION (WHO). WHO Coronavirus (COVID-19). Dashboard (Acessed on July 30, 2023). World Health Organization, 2023.

Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. JAMA Intern Med (2020) 180:934. doi: 10.1001/jamainternmed.2020.0994

Yan, R., Zhang, Y., Li, Y., Xia, L., Gou, Y., and Zhou, Q. (2020). Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science 360, 1444–1448.

Yao H, Song Y, Chen Y, Wu N, Xu J, Sun C, . . . Li S. Molecular Architecture of the SARSCoV-2 Virus. Cell. 183(3), 730-738.e13, 2020.

Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579, 270–273. doi: 10.1038/s41586-020-2012-7

6. COMPROVANTE DE SUBMISSÃO/ACEITE DE ARTIGO CIENTÍFICO

Journal of *Personalized Medicine*

an Open Access Journal by MDPI

Indexed in: PubMed | CITESCORE 1.8 | IMPACT FACTOR 3.508

CERTIFICATE OF PUBLICATION

Certificate of publication for the article titled:

The Genomic Profile Associated with Risk of Severe Forms of COVID-19 in Amazonian Native American Populations

Authored by:

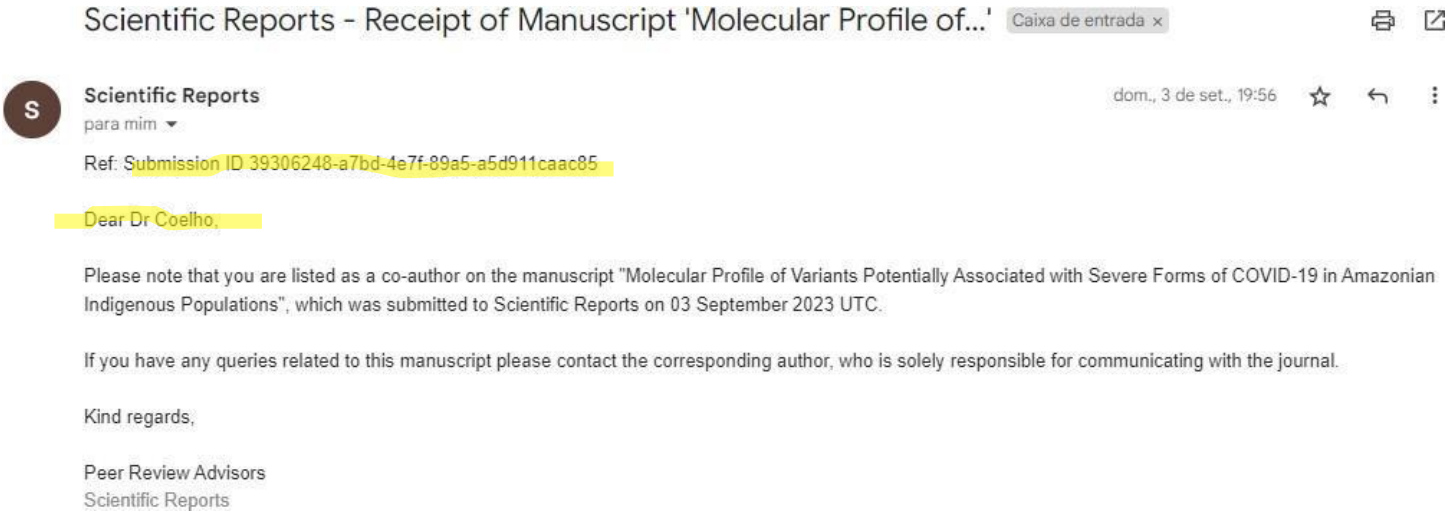
Lucas Favacho Pastana; Thays Amâncio Silva; Laura Patrícia Albarello Gellen; Giovana Miranda Vieira; Letícia Almeida de Assunção; Luciana Pereira Colares Leitão; Natasha Monte da Silva; Rita de Cássia Calderaro Coelho; Angélica Leite de Alcântara; Lui Wallacy Morikawa Souza Vinagre; Juliana Carla Gomes Rodrigues; Diana Feio da Veiga Borges Leal; Marianne Rodrigues Fernandes; Sandro José de Souza; José Eduardo Kroll; André Mauricio Ribeiro-dos-Santos; Rommel Mario Rodríguez Burbano; João Farias Guerreiro; Paulo Pimentel de Assumpção; Ândrea Campos Ribeiro-dos-Santos; Sidney Emanuel Batista dos Santos; Ney Pereira Carneiro dos Santos

MDPI Academic Open Access Publishing

Basel, October 2022

Published in:

J. Pers. Med. 2022, Volume 12, Issue 4, 554



Scientific Reports - Receipt of Manuscript 'Molecular Profile of...' Caixa de entrada ×

Scientific Reports
para mim

dom., 3 de set., 19:56

Ref: Submission ID 39306248-a7bd-4e7f-89a5-a5d911caac85

Dear Dr Coelho,

Please note that you are listed as a co-author on the manuscript "Molecular Profile of Variants Potentially Associated with Severe Forms of COVID-19 in Amazonian Indigenous Populations", which was submitted to Scientific Reports on 03 September 2023 UTC.

If you have any queries related to this manuscript please contact the corresponding author, who is solely responsible for communicating with the journal.

Kind regards,

Peer Review Advisors
Scientific Reports

APÊNDICE

Supplementary Table S1- Description of the 66 variants present in the genes AQP3, ARHGAP27, ELF5, IFNAR2, LIMD1, OAS1 and UPK1A.

Gene	Position	SNP ID	Ref[a]	Var[b]	Impact Predicted by SNPeff	Variant Allele Frequence					
						INDG	AFR	AMR	EAS	EUR	SAS
AQP3	33442954	rs2228332	G	A	LOW	0,945313	0.7462	0.7106	0.7177	0.5901	0.5756
AQP3	33447426	rs591810	C	G	LOW	0,96875	0.7678	0.8105	0.7486	0.703	0.6534
AQP3	33442274	rs2231235	G	A	MODIFIER	0	0.0126	0.1099	0.0633	0.0698	0.0859
AQP3	33442882	*	G	A	LOW	0,083333	-	-	-	-	-
AQP3	33443875	rs114247802	G	A	LOW	0,083333	0.0048	0.0007	-	-	-
AQP3	33442988	rs2231231	A	C	MODIFIER	0,9375	0.775	0.7272	0.7217	0.6398	0.5867
AQP3	33447549	rs12555686	G	T	MODIFIER	0	0.0221	0.255	0.1752	0.0951	0.0953
ARHGAP27	45430283	rs7222444	A	G	MODIFIER	0,009091	0.1027	0.158	0.0005	0.243	0.0717
ARHGAP27	45429642	rs2959953	G	C	MODERATE	0,375	0.6953	0.6537	0.5208	0.7177	0.6734
ARHGAP27	45395941	rs62064597	C	G	MODIFIER	0,083333	0.0424	0.132	0.0008	0.2141	0.0735
ARHGAP27	45396708	rs1297259327	C	T	LOW	0,083333	-	-	-	-	-
ARHGAP27	45430037	rs7220206	G	A	LOW	0,028571	0.2174	0.4237	-	0.4716	0.4716
ARHGAP27	45429583	rs115993362	C	G	MODIFIER	0,083333	0.1478	0.0062	-	0.0002	0.0006
ARHGAP27	45429931	rs12949256	C	T	MODERATE	0	-	-	-	0.5	0.1667
ARHGAP27	45410208	rs35327136	C	A	MODIFIER	0,015625	0.1308	0.1778	-	0.2233	0.0682
ARHGAP27	45396860	rs7213200	G	A	MODIFIER	0,016129	-	-	-	-	-
ARHGAP27	45410198	rs142163608	C	T	MODIFIER	0,083333	0.0162	-	-	-	-
ARHGAP27	45404234	rs184103721	C	G	MODIFIER	0,045455	0.0004	0.0481	0.0104	0.0002	0.0004
ARHGAP27	45410058	rs35389313	T	C	MODIFIER	0,013514	-	-	-	-	-
ARHGAP27	45404053	rs201721078	G	A	HIGH	0,21875	0.0009	0.0421	-	0.0001	-

Supplementary Table S1- Description of the 66 variants present in the genes AQP3, ARHGAP27, ELF5, IFNAR2, LIMD1, OAS1 and UPK1A.

ARHGAP27	45395459	rs117139057	G	C	MODERATE	0,125	0.089	0.2028	0.2768	0.1791	0.1189
ELF5	34489994	rs2231825	C	A	MODIFIER	0,046875	0.1168	0.1868	0.027	0.4292	0.3018
ELF5	34505605	rs2231821	A	G	MODIFIER	0,04386	0.0596	0.1556	0.0268	0.316	0.2659
ELF5	34480798	rs2231828	T	C	LOW	0,046875	0.0806	0.1829	0.027	0.4214	0.3137
ELF5	34493816	rs556840829	A	G	MODIFIER	0,05	-	-	-	-	-
ELF5	34511455	rs28395819	A	C	MODIFIER	0,083333	-	-	-	-	-
ELF5	34493394	rs737254	T	G	MODIFIER	0,040541	0.2868	0.3272	0.051	0.4674	0.3397
IFNAR2	33241950	rs1051393	T	G	MODERATE	0,789063	0.1931	0.4873	0.5838	0.3362	0.4762
IFNAR2	33241945	rs2229207	T	C	MODERATE	0,34375	0.0798	0.1663	0.1786	0.084	0.1302
IFNAR2	33262573	rs1131668	G	A	MODIFIER	0,795455	0.3768	0.4537	0.5833	0.3654	0.4767
IFNAR2	33262760	rs9984273	C	T	MODIFIER	1	0.5946	0.8484	0.9181	0.6873	0.7679
IFNAR2	33245424	rs2252639	A	T	MODIFIER	0,285714	1	1	1	1	1
IFNAR2	33262774	rs34865572 rs750263757	CT	C	LOW	0,137931	1	1	1	1	1
IFNAR2	33244908	rs2834158	T	C	MODIFIER	0,333333	0.8076	0.5129	0.4167	0.6638	0.5244
IFNAR2	33230489	rs17860118	G	T	MODIFIER	0,1875	0.0762	0.1189	0.1399	0.0864	0.1208
IFNAR2	33244934	rs149186597 rs79402470	TTT C	T	LOW	0,083333	0.0026	0.0116	0.0003	0.0254	0.0261
IFNAR2	33262748	rs3216172 rs397789038	G	GT	MODIFIER	0,181818	0.3665	0.467	0.5943	0.33	0.4847
IFNAR2	33263275	rs147568312	C	T	LOW	0,083333	0.0163	0.0004	0.0002	-	-
IFNAR2	33262799	*	C	A	MODERATE	0,083333	-	-	-	-	-
IFNAR2	33230629	rs12482556	T	C	MODIFIER	0,714286	-	-	-	-	-

Supplementary Table S1- Description of the 66 variants present in the genes AQP3, ARHGAP27, ELF5, IFNAR2, LIMD1, OAS1 and UPK1A.

IFNAR2	33252224	rs56197608	C	T	MODIFIER	0	-	0.0035	0.021	0.0002	0.0001
LIMD1	45595761	rs267237	C	T	MODERATE	0,9375	0.5847	0.7219	0.9384	0.5625	0.6755
LIMD1	45636143	rs11379880 rs5848761	C	CT	LOW	1	1	1	1	1	1
LIMD1	45595761	rs267237	C	T	LOW	0.624	0.5847	0.7219	0.9384	0.5625	0.6755
LIMD1	45595947	rs267236	T	C	LOW	0,945313	0.7101	0.7409	0.9385	0.6046	0.6991
LIMD1	45674275	rs2742409	G	A	MODIFIER	0,083333	-	-	-	-	-
LIMD1	45636143	rs11379880 rs5848761	C	CT	LOW	1	1	1	1	1	1
LIMD1	45676789	*	G	T	MODIFIER	0,066667	-	-	-	-	-
LIMD1	45674196	rs2578679	G	A	MODIFIER	0,5	-	-	-	-	-
LIMD1	45636146	rs5848762 rs869033908	CA	C	HIGH	1	1	1	1	1	1
LIMD1	45676835	rs7640969	T	G	MODIFIER	-	-	-	-	-	-
LIMD1	45636146	rs5848762 rs869033908	CA	C	LOW	1	1	1	1	1	1
OAS1	1,13E+08	rs7968145	C	T	MODIFIER	1	0.9891	0.955	1	0.9241	0.9552
OAS1	1,13E+08	rs1131476	G	A	MODIFIER	1	0.9346	0.8318	0.7843	0.6591	0.698
OAS1	1,13E+08	rs1131454	G	A	MODERATE	0,645161	0.2286	0.7033	0.6141	0.5752	0.572
OAS1	1,13E+08	rs2660	G	A	MODERATE	1	0.9346	0.8409	0.791	0.6794	0.7078
OAS1	1,13E+08	rs10774671	G	A	MODIFIER	0,833333	0.4246	0.7997	0.7837	0.6544	0.7017
OAS1	1,13E+08	rs7967461	G	C	MODIFIER	1	0.9177	0.7394	0.76	0.5752	0.692
OAS1	1,13E+08	rs11352835	TA	T	MODIFIER	1	0.9169	0.7387	0.7692	0.59	0.6916
OAS1	1,13E+08	rs7955146	C	T	LOW	0,013514	0.3015	0.0159	-	0.0015	0.0005

Supplementary Table S1- Description of the 66 variants present in the genes AQP3, ARHGAP27, ELF5, IFNAR2, LIMD1, OAS1 and UPK1A.

OAS1	-	rs1051042	G	C	MODIFIER	1	0.9346	0.8318	0.7841	0.6578	0.6933
UPK1A	35676195	rs747589460	CTT	C	MODIFIER	0,033333	0.0227	0.037	0.0417	0.0587	0.0811
			T								
UPK1A	35673206	rs75289222	G	T	MODIFIER	0,035714	0.0653	0.115	0.1029	0.1016	0.0845
UPK1A	35668466	rs2267586	T	G	MODERAT	0,179688	0.0715	0.1157	0.2241	0.1007	0.1026
					E						
UPK1A	35678012	rs2285421	T	C	MODIFIER	0,3	0.8012	0.5853	0.6216	0.5365	0.7368
UPK1A	35677970	rs2285420	G	A	LOW	0,166667	0.0694	0.1678	0.1273	0.1397	0.1087

[a]Reference Allele; [b]Variant Allele; *Variants without described SNP; (-) – No annotation; INDG: Indigenous populations; AFR: African populations; AMR: American populations; EAS: East Asian population; EUR: European population; SAS: South Asian population.

Supplementary Table S2 – Comparison of the allele frequency of the INDG population with the continental population (AFR, AMR, EUR, EAS and SAS).

Gene	SNP ID	INDG vs. AFR *	INDG vs. AMR *	INDG vs. EAS *	INDG vs. EUR *	INDG vs. SAS *
AQP3	rs2228332	0.00018	3.693×10^{-5}	5.348×10^{-5}	4.293×10^{-9}	9.082×10^{-10}
AQP3	rs591810	3.963×10^{-5}	0.00075	1.031×10^{-5}	3.214×10^{-7}	1.087×10^{-8}
AQP3	rs2231235	0.09707	0.00324	0.42487	5.369×10^{-11}	1.929×10^{-6}
AQP3	rs114247802	0.00019	7.994×10^{-5}	NA	NA	NA
AQP3	rs2231231	0.00117	0.00010	5.442×10^{-5}	1.620×10^{-7}	2.264×10^{-9}
AQP3	rs12555686	0.63286	6.573×10^{-8}	2.387×10^{-5}	0.00362	0.00356
AQP3	rs16919255	0.26429	0.00027	1.570×10^{-5}	0.14649	0.14583
ELF5	rs2231825	0.09707	0.00324	0.42487	5.369×10^{-11}	1.929×10^{-6}
ELF5	rs2231821	1.000	0.01790	0.42487	6.224×10^{-7}	3.250×10^{-5}
ELF5	rs2231828	0.46491	0.00498	0.42487	1.168×10^{-10}	1.051×10^{-6}
ELF5	rs737254	4.787×10^{-6}	6.137×10^{-7}	1.000	1.465×10^{-12}	1.611×10^{-7}
UPK1A	rs747589460	0.65377	1.000	1.000	0.56369	0.20915
UPK1A	rs75289222	0.41762	0.04259	0.07032	0.07073	0.21094
UPK1A	rs2267586	0.00344	0.14912	0.63094	0.05448	0.05570
UPK1A	rs2285421	3.016×10^{-16}	3.180×10^{-5}	1.341×10^{-6}	0.00030	1.562×10^{-11}
UPK1A	rs2285420	0.01134	1.000	0.32717	0.45227	0.14569
OAS1	rs7968145	1.000	0.14897	1.000	0.01497	0.09495
OAS1	rs1131476	0.02583	4.458×10^{-5}	7.665×10^{-7}	3.188×10^{-11}	6.497×10^{-10}
OAS1	rs1131454	3.523×10^{-11}	0.37592	0.78508	0.34783	0.34646
OAS1	rs2660	0.02583	7.618×10^{-5}	1.374×10^{-6}	1.525×10^{-10}	2.482×10^{-9}
OAS1	rs10774671	3.039×10^{-10}	0.73237	0.51581	0.00451	0.03860
OAS1	rs7967461	0.01055	3.650×10^{-8}	1.260×10^{-7}	1.118×10^{-14}	6.043×10^{-10}
OAS1	rs11352835	0.01087	3.432×10^{-8}	2.366×10^{-7}	3.623×10^{-14}	6.043×10^{-10}

OAS1	rs7955146	1.422×10^{-8}	1.000	NA	0.21318	0.11573
OAS1	rs1051042	0.02583	4.458×10^{-5}	7.665×10^{-7}	3.220×10^{-11}	5.950×10^{-10}
LIMD1	rs267237	1.489×10^{-9}	6.326×10^{-5}	1.000	3.942×10^{-10}	3.273×10^{-6}
LIMD1	rs267236	2.363×10^{-5}	0.00028	1.000	9.931×10^{-9}	1.806×10^{-5}
ARHGAP27	rs7222444	0.02289	0.00064	0.11267	1.540×10^{-6}	0.10633
ARHGAP27	rs2959953	6.052×10^{-7}	4.116×10^{-5}	0.03356	1.238×10^{-7}	6.31360×10^{-6}
ARHGAP27	rs62064597	0.19781	0.30226	1.57513×10^{-5}	0.00765	0.80278
ARHGAP27	rs7220206	0.00011	2.062×10^{-11}	NA	$7.884 \text{ X } 10^{-14}$	$2.465 \text{ X } 10^{-11}$
ARHGAP27	rs115993362	0.13752	0.00129	NA	1.589×10^{-5}	1.801×10^{-5}
ARHGAP27	rs35327136	0.00401	0.00022	NA	795.025	0.16105
ARHGAP27	rs142163608	0.00890	NA	NA	NA	$1.801 \text{ X } 10^{-5}$
ARHGAP27	rs184103721	0.00063	1.000	0.05055	0.00137	0.00148
ARHGAP27	rs201721078	5.129×10^{-15}	1.542×10^{-5}	NA	$1.381 \text{ X } 10^{-14}$	NA
ARHGAP27	rs117139057	0.36469	0.16845	0.00951	0.37949	0.83868
IFNAR2	rs1051393	1.824×10^{-21}	1.053×10^{-5}	0.00258	1.215×10^{-11}	3.942×10^{-6}
IFNAR2	rs2229207	2.456×10^{-8}	0.00186	0.003	1.035×10^{-7}	6.205×10^{-5}
IFNAR2	rs1131668	1.229×10^{-10}	2.771×10^{-7}	0.00098	7.147×10^{-11}	9.955 x 10-7
IFNAR2	rs9984273	3.001×10^{-14}	0.00013	0.00938	3.134×10^{-10}	2.268×10^{-7}
IFNAR2	rs2834158	5.32×10^{-15}	0.00664	0.22382	3.818×10^{-7}	0.00342
IFNAR2	rs17860118	0.00771	0.15350	0.34659	0.02174	0.16154
IFNAR2	rs149186597 rs79402470	8.004×10^{-5}	0.00602	1.575×10^{-5}	0.04188	0.04602
IFNAR2	rs3216172 rs397789038	0.00380	2.699×10^{-5}	4.368×10^{-10}	0.02185	3.997×10^{-6}

IFNAR2	rs147568312	0.00890	7.994×10^{-5}	1.575×10^{-5}	NA	NA
IFNAR2	rs56197608	NA	1.000	0.62213	1.000	1.000

INDG: Indigenous populations; AFR: African populations; AMR: American populations; EAS: East Asian population; EUR: European population; SAS: South Asian population; *p-value defined by Fisher's exact test. Bold characters indicate a significant difference (p-value* < 0.05).