UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

MARIA CLARA DA COSTA BARROS

CARACTERIZAÇÃO DO EXOMA DE POPULAÇÕES HUMANAS DA AMAZÔNIA EM RELAÇÃO À INFECÇÃO PELO *SARS-CoV-2*.

BELÉM – PA

Fevereiro de 2023

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

MARIA CLARA DA COSTA BARROS

CARACTERIZAÇÃO DO EXOMA DE POPULAÇÕES HUMANAS DA AMAZÔNIA EM RELAÇÃO À INFECÇÃO PELO *SARS-CoV-2*.

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Federal do Pará, como requisito final para obtenção do título de mestre.

Orientador: Prof. Dr. Sandro José de Souza.

BELÉM – PA

Fevereiro de 2023

AGRADECIMENTOS

À minha avó Tereza Costa. A saudade é imensurável e mesmo sabendo que a senhora continua acompanhando e apoiando os meus passos, eu ainda queria que a senhora estivesse aqui presencialmente para dividirmos essa vitória juntas.

À minha família por todo apoio, amor e incentivo incondicionais. Por sempre acreditar em mim e nunca duvidar do meu potencial. Um agradecimento especial aos meus pais, meu irmão e meus avós, por fazer o possível e impossível pela minha educação tanto acadêmica quanto de vida. Vocês são o maior e melhor exemplo de perseverança, devoção, amor e resiliência que eu poderia ter.

Ao meu cachorro, Lord, por me ensinar o amor mais puro e sincero que não espera nada em troca.

À ciência que me encanta constantemente por ser mutável e imutável, por ser a base da evolução da humanidade e por mesmo que sejamos feitos “apenas” de poeira de estrelas (obrigada Carl Sagan), somos indivíduos singulares com ambições e o poder de mudar a história. Creio que nunca fui tão fascinada pela ciência quanto fui e sou durante esses dois anos.

Durante esses dois anos fui agraciada de ter orientadores no plural, então primeiramente ao meu orientador Sandro de Souza, por me aceitar como aluna de mestrado e me acompanhar no seu laboratório em Natal.

À educadora Ândrea Ribeiro-dos-Santos por quem eu tenho uma imensa admiração desde a graduação e sou imensamente grata por ter me aceitado como aluna na pós. Obrigada por ser exemplar e por me permitir crescer como acadêmica, profissional e pesquisadora e por todas as oportunidades conquistadas.

Ao professor Jorge por todo o conhecimento compartilhado, por toda a disponibilidade e por toda a paciência ao me ensinar sobre o mundo complicado e encantador repleto de possibilidades que é a bioinformática.

À Natal, por ser uma cidade deslumbrante e tão familiar e me proporcionar uma das melhores experiências da minha vida que eu nunca esquecerei. Ao BioMe e as pessoas que fazem parte, Bianca Santiago, Matheus Miguel e Epitácio Dantas, obrigada por todo acolhimento.

À Rafaella Ferraz, Giovanna Cavalcante e Leandro Magalhães. Sou eternamente grata pelo apoio e paciência e por todos os ensinamentos que irei levar para a vida pessoal e profissional. Eu não seria a pesquisadora que sou sem vocês, vocês se tornaram minha inspiração e meu exemplo nessa vida árdua e extraordinária da pesquisa.

À equipe do LabCoV: Catarina Pinho, Caio Santos, Cíntia Silva, Jhully Azevedo e Laura Closset. Por toda a pesquisa linda que nós realizamos e ter ajudado tantas pessoas durante esses dois anos de pandemia. O exoma não seria realizado sem vocês e sou extremamente agradecida por isso.

À todos aqueles não citados aqui mas que não deixam de serem importantes na minha formação.

E por fim, à Universidade Federal do Pará por todas as oportunidades e momentos incríveis vivenciados, nossa relação é de amor e ódio com altos e baixos, mas continuo sendo eternamente grata por tudo.

SUMÁRIO

1. INTRODUÇÃO 9
1.1. COVID-19 9
1.2. SEQUENCIAMENTO DE NOVA GERAÇÃO 14
1.2.1. EXOMA 16
1.3. ANCESTRALIDADE GENÔMICA E COVID-19 20
1.4. BIOINFORMÁTICA 21
2. JUSTIFICATIVA 24
3. OBJETIVOS 26
3.1 OBJETIVO GERAL 26
3.2. OBJETIVOS ESPECÍFICOS 26
4. METODOLOGIA 27
4.1. COMITÊ DE ÉTICA 28
4.2. CARACTERIZAÇÃO DAS AMOSTRAS 28
4.3. ANÁLISE LABORATORIAL 30
4.3.1 Extração de DNA 30
4.3.2 Preparação da biblioteca do exoma 30
4.4 ANÁLISE BIOINFORMÁTICA 32
4.4.1. Programas e ferramentas 32
4.4.1.1. Trimmomatic 32
4.4.1.2. BWA MEM 32
4.4.1.3. SAMtools 32
4.4.1.4. PICARD Tools 33
4.4.1.5. GATK 33
4.4.1.6. SnpEff 33
4.4.2. Bancos de referências 34
4.4.2.1. ClinVar 34
4.4.2.2. COSMIC 34
4.4.2.3. dbSNP 34
4.4.2.4. BRCA1 34
4.4.2.5. ExAC 35
4.4.3. Processamento de reads e chamada de variantes 35
4.5. ANÁLISES GENÔMICO-POPULACIONAIS 38

4.6. ANÁLISES ESTATÍSTICAS 39
5. RESULTADOS 40
5.1 ANÁLISE CLÍNICA 40
5.2 ANÁLISE DESCRITIVA DAS VARIANTES 42
5.3 ANÁLISE DA ANCESTRALIDADE GENÔMICA 51
5.4 ANÁLISE DE ASSOCIAÇÃO GENÉTICA 54
6. DISCUSSÃO 57
7. CONCLUSÃO 68
REFERÊNCIAS 70

RESUMO

A COVID-19 causada pelo agente etiológico SARS-CoV-2 é considerada uma das pandemias mais graves na história humana moderna. Até o presente momento (dia 4 de fevereiro de 2023), foram confirmados 754.018.841 casos e 6.817.478 mortes por COVID-19 no mundo, sem contar os casos de subnotificação. A COVID-19 difere das demais doenças pelo seu amplo espectro clínico, englobando desde casos assintomáticos ou leves até casos graves de insuficiência respiratória e óbito. Esta amplitude de sintomas clínicos pode ser decorrente das diferentes linhagens virais, mas também pode ocorrer entre indivíduos infectados pela mesma linhagem, sugerindo que há outros fatores envolvidos no agravamento da doença, além de fatores socioeconômicos e das diferentes políticas de combate à COVID-19 adotadas por cada país. Com o objetivo de caracterizar marcadores genéticos associados à gravidade clínica da infecção por *SARS-CoV-2*, foi realizado o sequenciamento total do exoma de indivíduos que apresentaram tanto a forma não grave, quanto a mais grave no espectro dessa doença. Em nosso estudo, assim como em outros, também foi encontrado associação entre a idade (*pvalor* = 0.005; IC 95% = 8.98 – 21.2), linhagem viral (*pvalor* = 0.05) e presença de comorbidades (*pvalor* = 0.01; OR= 9.3; IC 95% = 2.97 – 33.05) a gravidade clínica da COVID-19. Adicionalmente, nossos resultados descreveram cinco variantes genômicas, ainda não identificadas em estudos prévios, que podem estar associadas a COVID-19 e a gravidade clínica da doença. Essas variantes estão presentes nos genes *MYLK* (rs40305), *ERAP1* (rs469783), *PRAG1* (rs4840953), *MCM3AP* (rs2839166) and *NFIB* (rs12349255). Até onde sabemos, esse é o primeiro estudo de sequenciamento total do exoma em indivíduos brasileiros, com ênfase na região Amazônica, que busca investigar a parte genética dos mecanismos associados à patogênese da gravidade clínica da COVID-19, reforçando a importância de se realizar estudos genômicos mais representativos das distintas populações brasileiras.

Palavras-chave: Exoma; *SARS-CoV-2*; Gravidade Clínica; Marcadores Genômicos; População Amazônica.

ABSTRACT

COVID-19 caused by the etiological agent identified as SARS-CoV-2 is considered one of the most serious pandemics in modern history. To date (February 4, 2023), 754,018,841 cases and 6,817,478 deaths from COVID-19 have been confirmed worldwide, not counting underreported cases. COVID-19 differs from other diseases due to its broad clinical spectrum, ranging from asymptomatic or mild cases to severe cases of respiratory failure and death. This amplitude in clinical symptoms may be due to different viral strains, but it can also occur in individuals infected by the same strain, suggesting that there are other factors involved in the worsening of the disease, in addition to socioeconomic factors and the different policies adopted to combat COVID-19 for each country. With the aim of characterizing genetic markers associated with the clinical severity of SARS-CoV-2 infection, whole exome sequencing of individuals who presented the non-severe and severe forms of this disease was performed. An association was found between age (pvalue =0.005; CI 95% = 8.98 – 21.2), viral lineage (pvalue = 0.05), the presence of comorbidities (pvalue = 0.01; OR= 9.3; CI 95% = 2.97 – 33.05) and the clinical severity of COVID-19. Five variants possibly associated with the clinical severity of COVID-19 were observed in this study, however none have previously been associated with COVID-19. These variants are present in the genes: *MYLK* (rs40305), *ERAP1* (rs469783), *PRAG1* (rs4840953), *MCM3AP* (rs2839166) and *NFIB* (rs12349255). To the best of our knowledge, this is the first whole exome sequencing study in Brazilian individuals, with emphasis in the Amazon Region, that seeks to investigate the genetic part of the mechanisms associated with the pathogenesis of the clinical severity of COVID-19, reinforcing the importance of carrying out more genomic studies in Brazilian distinct populations.

Key-words: Exome; *SARS-CoV-2*; Clinical Severity; Genomic markers; Amazonian population.

1. INTRODUÇÃO

1.1. COVID-19

A COVID-19 causada pelo agente etiológico identificado como *SARS-CoV-2* (Síndrome Respiratória Aguda Grave do Coronavírus 2), é considerada uma das pandemias mais graves na história moderna (AMOROSO et al., 2021; KHANNA et al., 2020; ZHOU et al., 2020). Oficialmente surgiu em dezembro de 2019, quando casos de pneumonia de origem desconhecida foram reportados em Wuhan, província de Hubei na China. Em março de 2020, a COVID-19 foi declarada como pandemia pela Organização Mundial da Saúde devido sua elevada taxa de disseminação em mais de cem países, apenas dois meses após a sua descoberta (KHANNA et al., 2020; ONYEAKA et al., 2021; RUIZ-STERNBERG et al., 2022). Até o presente momento (dia 1 de setembro de 2022), foram confirmados 754.018.841 casos e 6.817.478 mortes por COVID-19 no mundo, sem contar os casos de subnotificação (WHO, 2022).

O primeiro caso de COVID-19 no Brasil e, também da América do Sul foi detectado em um homem de 61 anos que havia viajado para a Itália, sendo reportado no dia 26 de fevereiro de 2020 (ARAÚJO et al., 2020; XAVIER et al., 2020). Dois anos e meio depois já foram reportados mais de 36 milhões de casos e 697.000 óbitos aproximadamente, resultando no colapso do sistema de saúde brasileiro tanto público quanto particular (WHO, 2023).

Na região Norte do Brasil, o primeiro caso oficial de COVID-19 foi descrito no dia 13 de março de 2020 na cidade de Manaus (PEREIRA et al., 2020). Em oposição a alguns estudos iniciais que sugeriam que o vírus se tornava inativado em países com

temperaturas elevadas, os primeiros casos identificados foram em onze regiões de clima quente ao longo do rio Amazonas, com prevalência mais elevada do que o esperado e abrangendo áreas indígenas (HALLAL et al., 2020). Ao todo, foram contabilizados aproximadamente 2.745.000 casos e 50 mil óbitos na Região Norte, principalmente nos estados do Pará e Amazonas (MINISTÉRIO DA SAÚDE, 2023).

Ademais, foi possível observar que a distribuição do vírus não tem sido homogênea, ocasionando diferenças significativas nas taxas de transmissão e de morte entre países e continentes. De janeiro de 2020 a setembro de 2022, por exemplo, os países do continente americano apresentaram elevadas taxas de mortalidade (acima de 300 mil indivíduos), enquanto os países africanos tinham menos de mil a oito mil óbitos (WHO, 2022). Sabe-se que a taxa de letalidade é influenciada pela velocidade de contágio e, principalmente, por peculiaridades das populações afetadas, incluindo as diferentes políticas de contenção adotadas, condições socioeconômicas e o apoio do sistema de saúde característico de cada população (JIN et al., 2020; YANG et al., 2020). É importante ressaltar que fatores sociais e econômicos possuem uma grande influência na gravidade da COVID-19 de cada indivíduo, entretanto não podemos afirmar que sejam os únicos a influenciar o desfecho dessa doença.

Sendo assim, uma das estratégias para melhor controle e tratamento da doença consiste na melhor compreensão dos mecanismos de ação do vírus. Com esse propósito, o agente etiológico da COVID-19 foi identificado após a análise metagenômica utilizando o Sequenciamento do Genoma Completo realizado nos primeiros pacientes infectados, sendo esse procedimento essencial para a exclusão da possibilidade de ser apenas uma pneumonia, como haviam descrito inicialmente (ZHOU et al., 2020).

Do ponto de vista biológico, o *SARS-CoV-2* é uma partícula envelopada e esférica de 120 nm de diâmetro constituída por um RNA de fita positiva simples com

aproximadamente 29.9 mil pares de base, além da cap 5’ metil-guanina e a cauda poli-A, o que permite a sua tradução a partir da maquinaria de tradução do hospedeiro (KHANNA et al., 2020; LIU; KUO; SHIH, 2020; MARIANO et al., 2020). Esse vírus pertence à família *Coronaviridae*, do gênero *Betacoronavírus* (linhagem B), sendo o sétimo dessa família a infectar humanos, após o *HCoV-OC43, SARS-CoV, MERS-CoV*, entre outros (YIN; WUNDERINK, 2018; CHAN et al., 2020; YAO et al., 2021).

Alguns dos genomas virais sequenciados em diversos países foram geneticamente (BAMSHAD et al., 2011) idênticos, enquanto outros carregavam mutações distintas, destacando a capacidade do vírus de mutar e se ambientar às diferentes populações humanas e distintos ambientes ecológicos (KHARE et al., 2021).

Ademais, a COVID-19 difere das demais doenças pelo seu amplo espectro clínico, englobando desde casos assintomáticos ou leves (sintomas gripais), até casos graves de insuficiência respiratória e óbito. Esta variação de sintomas clínicos pode ser decorrente das diferentes linhagens virais, mas também pode ocorrer em indivíduos infectados pela mesma linhagem, sugerindo que há outros fatores envolvidos no agravamento da doença, além de fatores socioeconômicos e das diferentes políticas de combate à COVID-19 adotadas por cada país (IZCOVICH et al., 2020; BALDASSARRI et al., 2021; ANDOLFO et al., 2021).

A patogênese do SARS-COV-2 inicia-se pela disseminação de gotículas no trato respiratório, principalmente nas células epiteliais alveolares, transmitidas de indivíduo para indivíduo (RAHMAN et al., 2021; ROVITO et al., 2022). Em seguida, a entrada do vírus nas células do hospedeiro ocorre por meio da ligação do domínio de ligação ao receptor (RBD) na subunidade S1 da glicoproteína Spike, presente na superfície do *SARS-CoV-2*, ao receptor da superfície da célula hospedeira, principalmente a enzima

conversora de angiotensina 2 (ACE2) (TAN & KOMARASAMY & BALASUBRAMANIAM, 2021; RAVI & SAXENA & PANDA, 2022).

A ligação da proteína Spike com a proteína ACE2 induz alterações conformacionais na subunidade S1, expondo a subunidade S2 da proteína. A partir disso, a subunidade S2 pode ser clivada de duas formas: endocitose mediada por clatrinas e fusão direta do envelope viral com a membrana celular. O primeiro processo ocorre se a célula-alvo hospedeira não expressa a protease transmembrana serina 2 (TMPRSS2) de forma suficiente ou não há essa proteína, então o complexo vírus-ACE2 é internalizado via endocitose mediada por clatrinas no endolisossomo, onde a subunidade S2 é clivada por catepisinas. Caso haja TMPRSS2 suficiente na membrana celular, a própria facilita a fusão do vírus na célula, uma vez que cliva a subunidade S2 e libera o RNA viral no citoplasma (RAHMAN et al., 2021; LEE & KIM & AHN, 2022; JACKSON et al., 2022).

Uma vez no citoplasma, o RNA dos coronavírus pode produzir diretamente suas proteínas e novos genomas ao entrar nos ribossomos do hospedeiro, que por sua vez traduzem o RNA viral para RNA-polimerases e posteriormente em fitas simples positivas de RNA. A partir disso, ocorre a formação dos componentes estruturais do vírus no retículo endoplasmático e depois a sua montagem final no complexo de Golgi para enfim serem liberados da célula hospedeira por exocitose por meio de vesículas secretoras (RAHMAN et al., 2021; BOSON et al., 2021; LEE & KIM & AHN, 2022).

Apesar disso, os fatores patogênicos específicos para um indivíduo desenvolver a forma grave da doença ainda permanecem incertos. Entretanto, estudos já demonstraram que uma resposta inflamatória eficiente permite uma evolução benigna da COVID-19, sem sequelas, enquanto a forma grave ocorre em pacientes que apresentam uma resposta imune exacerbada ao *SARS-CoV-2* (CHEN et al., 2020; MERAD et al., 2022; ROVITO et al., 2022).

À vista disso, após infectar a célula hospedeira, o *SARS-CoV-2* é reconhecido nos endossomos por receptores de reconhecimento padrão (PRRs), como os receptores Toll-like (TLR 3, TLR7 e TLR8) que reconhecem a fita simples positiva de RNA e recrutam interferons (IFNs) tipo 1 e tipo 3, além do complexo proteico NF-κB (fator nuclear kappa B), que contribuem para a produção de quimiocinas e citocinas (FITZGERALD & KAGAN, 2020; BRODIN, 2021; MERAD et al., 2022).

Em relação a isso, já foi observada a "tempestade de citocinas", induzida pela imunidade inata, em casos graves de pacientes infectados pelo SARS-CoV-2, sendo os monócitos, macrófagos, neutrófilos, células NK e células dendríticas as principais células imunes envolvidas nesse processo, uma vez que produzem as citocinas e quimiocinas pró-inflamatórias: IL-1β, IL-2, IL-6, IL-7, IL-10, TNF-α e IFN-γ (WANG et al., 2020b; LEE & KIM & AHN, 2022). Portanto, indivíduos que apresentam alguma inflamação no organismo por conta de fatores de risco já associados a essa patologia, possuem uma predisposição maior a desenvolver a forma grave, uma vez que a tempestade de citocinas é agravada.

Sendo assim, os fatores de risco do hospedeiro mais comuns da COVID-19 são: idade avançada, sexo masculino, comorbidades prévias como hipertensão, diabetes, doenças cardiovasculares e até mesmo tipo sanguíneo (IZCOVICH et al., 2020; ZHENG et al., 2020; D'ANTONIO et al., 2021; ANGULO-AGUADO et al., 2022). Todavia, esses fatores de risco não explicam o amplo espectro na variação das manifestações clínicas da doença, visto que não foram somente os indivíduos pertencentes aos grupos supracitados que desenvolveram a forma grave da doença, sendo também observado o desenvolvimento desta, caracterizada pela saturação baixa de oxigênio e respostas inflamatórias no pulmão, em indivíduos jovens e/ou que não apresentaram comorbidades prévias (WANG et al., 2020c; OCHANI et al., 2021; O'DRISCOLL et al., 2021;

BRODIN, 2021). Estas características sugerem que a complexidade da interação vírus – hospedeiro não depende apenas de fatores socioeconômicos ou ambientais.

1.2. SEQUENCIAMENTO DE NOVA GERAÇÃO

Diante dessa perspectiva, uma das estratégias é investigar a interação vírus-hospedeiro do ponto de vista genético por meio do sequenciamento genômico massivo paralelo que faz parte do aparato tecnológico do Sequenciamento de Nova Geração (NGS), juntamente com as plataformas de estudos de GWAS amplamente utilizadas para identificar fatores genéticos do hospedeiro envolvidos na susceptibilidade a doenças e na farmacogenômica. (DEHGHAN, 2018; PANDIT et al., 2022).

Recentemente os Estudos de Associação Genômica Ampla (GWAS) em larga escala apresentaram fatores genéticos do hospedeiro que podem estar associados tanto à suscetibilidade do *SARS-CoV-2* quanto à gravidade da COVID-19 (ANASTASSOPOULOU et al., 2020; THE SEVERE COVID-19 GWAS GROUP, 2020; COVID-19 HOST GENETICS INITIATIVE et al., 2021; DOWNES et al., 2021; PAIRO-CASTINEIRA et al., 2021; ANGULO-AGUADO et al., 2022).

Do ponto de vista da susceptibilidade, alguns genes *ACE*, *ACE2*, *TMPRSS*, *FURIN* e interleucinas, como a *IL-6*, têm sido proposto como candidatos, uma vez que já possuem elevada afinidade com a proteína S do *SARS-CoV-2*, assim como contribuem para a rápida disseminação do vírus no organismo por facilitar sua entrada na célula (HUANG & WANG, 2021; YILDIRIM et al., 2021; ISHAK et al., 2022). Além disso, a sua expressão (elevada ou baixa) nos indivíduos, influencia na maior suscetibilidade à

infecção por *SARS-CoV*-2 (DEVAUX; ROLAIN; RAOULT, 2020; HOFFMANN et al., 2020; CHOUDHARY et al., 2021).

No que diz respeito à gravidade da clínica, que é o foco do presente estudo, genes que regulam a imunidade ao interferon tipo 1 como *TLR7*, *ZC3HAV1* e *IFNAR1* também já foram associados (PLENGE, 2020; ZHANG et al., 2020; KOSMICKI et al., 2021). Além deles, polimorfismos do sistema de antígeno leucocitário humano (HLA) também podem influenciar a resposta imune do indivíduo, uma vez que a quantidade de sítios de ligação ao *SARS-CoV*-2 varia entre os seus alelos (NGUYEN et al., 2020).

Entretanto, o principal achado genético associado ao quadro de agravamento, com a hospitalização de pacientes com COVID-19, foi o locus gênico presente na região 3p21.31 que confere um risco duas vezes maior de insuficiência respiratória por COVID-19 e um risco duas vezes maior de mortalidade para indivíduos com menos de 60 anos (THE SEVERE COVID-19 GWAS GROUP, 2020; NAKANISHI et al., 2021; SHELTON et al., 2021). Presentes nesse haplótipo, herdado dos Neandertais, estão os genes correceptores de *SARS-CoV-2 SLC6A20*, *LZTFL1* e *FYCO1* e os receptores de quimiocinas *CCR9*, *CXCR6* e *XCR1* (ZEBERG; PÄÄBO, 2020; DOWNES et al., 2021; KHALIL; ELEMAM; MAGHAZACHI, 2021; PAIRO-CASTINEIRA et al., 2021; YAO et al., 2021).

Todavia, como grande parte dos estudos são realizados em indivíduos europeus ou asiáticos, os achados não se tornam representativos ou aplicáveis às populações pouco investigadas ou miscigenadas. Até o momento, estudos buscando variantes genéticas do hospedeiro associadas a COVID-19 em indivíduos da América do Sul são bastante limitados.(MALAQUIAS et al., 2021; ANGULO-AGUADO et al., 2022).

No Brasil, apenas um GWAS foi realizado (BRACOVID). Esse estudo confirmou a maioria dos loci já identificados na literatura e identificou um novo loci associado à

hospitalização por COVID-19 em cinco mil indivíduos residentes na região Sudeste (PEREIRA et al., 2022). Entretanto, os GWAs realizados em doenças poligênicas (influenciada por muitos genes), como a COVID-19, além de identificarem apenas regiões já descritas na literatura, a chance de encontrar associações significativas são baixas visto que a análise é realizada em todo o genoma e não somente nas regiões codificantes e de alto impacto funcional (TAM et al., 2019).

1.2.1. EXOMA

Portanto, outra forma de se investigar variantes genéticas associadas às doenças é por meio do Sequenciamento Completo do Exoma (WES). O exoma refere-se às regiões codificantes do genoma, compreendendo 180.000 exóns em aproximadamente 22.000 genes (BAMSHAD et al., 2011). A partir do seu sequenciamento, é possível analisar e interpretar as alterações em todas as regiões codificantes de proteínas de um indivíduo, sejam genes previamente conhecidos ou não, além de regiões intrônicas adjacentes (NG et al., 2010; LADUCA et al., 2017).

Além disso, o WES possui uma abordagem eficaz na análise genômica, pois além de ser uma técnica de custo mais baixo e requerer menor quantidade de material biológico, quando comparada ao GWAS por exemplo, fornece uma alta cobertura para identificação de variantes raras (de baixa frequência) (GUERRA et al., 2019). Dito isso, como o WES busca variantes e mudanças estruturais como CNVs (copy number variation), inserções e deleções, é possível identificar com mais clareza as possíveis causas de como os genes podem influenciar no agravamento ou susceptibilidade da doença em questão (BEALS, 2010; ELLINGFORD et al., 2017).

As principais etapas do exoma são: preparação da biblioteca, onde há a tagmentação, captura por hibridação, amplificação por PCR das sequências de DNA e quantificação do produto final para verificar a adequação da amostra para posterior sequenciamento massivo genômico e pôr fim a análise bioinformática dos dados gerados (REHM et al., 2013; HEAD et al., 2014).

A primeira aplicação do WES foi em doenças hereditárias, ou seja, em famílias em que haviam múltiplos indivíduos afetados a fim de encontrar variantes observadas apenas nesses casos e não em parentes não afetados (GILISSEN et al., 2012). Portanto, uma das vantagens da análise do exoma é poder determinar se a variante causadora da desordem genética foi herdada dos pais ou gerada espontaneamente (mutação nova) durante a gametogênese (PETROVSKI et al., 2019). Esta análise é a principal característica dos diversos estudos que empregam WES no período pré-natal para detecção de anormalidades fetais e no acompanhamento pós-natal após uma observação fenotípica específica de alguma manifestação clínica (LEI et al., 2017; MENG et al., 2017; CHANDLER et al., 2018; ROSS; DION; ROULEAU, 2020).

Além disso, a metodologia do WES pode ser utilizada tanto no diagnóstico (Sequenciamento Completo de Exoma de diagnóstico) como na descoberta de novas variantes (Sequenciamento Completo de Exoma de Pesquisa) em doenças com grande variabilidade genética, justamente por poder analisar um número significativamente maior de genes, permitir a avaliação direta de variantes envolvidas na alteração de proteínas, identificar novos genes e sequenciar exomas de vários indivíduos simultaneamente (FARWELL et al., 2015; LADUCA et al., 2017; VAN HOUT et al., 2019).

Em relação à susceptibilidade e gravidade da COVID-19, é possível analisar o exoma de pacientes acometidos, em busca de identificar variantes genéticas que possam

estar associadas, como por exemplo: variantes presentes em um gene específico (BALDASSARRI et al., 2021), variantes patogênicas raras (CANTALUPO et al., 2021; KOSMICKI et al., 2021), análise de enriquecimento em variantes com perda de função (VAN DER MADE et al., 2020; ZHANG et al., 2020), e *loci* gênico em populações específicas (CHAMNANPHON et al., 2022).

No Brasil, há poucos exomas de indivíduos infectados pela COVID-19 até o presente momento. Castelli e colaboradores (2021) analisaram o exoma de 83 casais brasileiros da Região Sudeste, onde um indivíduo estava infectado e sintomático e o parceiro permaneceu assintomático e soro-negativo por pelo menos seis meses mesmo compartilhando o mesmo ambiente doméstico, a fim de identificar mecanismos de resistência a infecção pelo vírus presentes na região do complexo principal de histocompatibilidade (MHC). Foi demonstrado uma associação entre alelos do gene *HLA* e susceptibilidade a infecção sintomática, corroborando a importância de estudar genes associados a modulação imune (CASTELLI et al., 2021).

Além desse estudo, Khayat e colaboradores (2021) investigaram a presença de variantes no gene *ACE2*, em 64 ameríndios e 82 indivíduos residentes na Região Norte e compararam com as frequências alélicas observadas nas populações do banco de dados *1,000 Genomes* (FAIRLEY et al., 2020). Foi observado que as populações ameríndias apresentavam mutações exclusivas e/ou cuja frequência era maior do que em outras populações, associados a uma maior expressão do gene *ACE2* em tecidos neurais (KHAYAT et al., 2020), entre outros. Estes resultados se mostraram um achado de grande relevância uma vez que poderia ter auxiliado no manejo e desfecho da doença nessas populações, quando contaminada pelo vírus (COVID-19 E OS POVOS INDÍGENAS, 2022).

Em 2021, Secolin e colaboradores analisaram 27 genes associados a COVID-19 em dois bancos de dados públicos brasileiros (257 indivíduos do BIPMed e 609 indivíduos do ABraOM) e em 88 exomas de indivíduos da Região Sudeste. No total, foram encontradas mais de sete mil variantes, dentre as quais 70 foram exclusivas da população brasileira, incluindo variantes em: a) genes relacionados à imunidade do interferon tipo I (INF) ao vírus influenza; b) genes candidatos à resposta à COVID-19; e c) variantes relacionadas a entrada do vírus SARS-*CoV-2* nas células pulmonares e sua replicação (SECOLIN et al., 2021). Apesar da relevância dos achados, a maior parte desses estudos se restringiu a avaliação de poucos genes nas populações selecionadas.

Portanto, a aplicação dos resultados obtidos pelo WES na clínica é impactante (SCHWARZE et al., 2018), pois permite uma melhor compreensão sobre os diferentes fatores genéticos e a progressão da infecção por COVID-19, crucial para o desenvolvimento de novas opções terapêuticas, estratégias de prevenção e diagnóstico precoce, que pode evitar o surgimento de pacientes clinicamente graves por meio de intervenção médica.

Além disso, a gravidade clínica aparenta ser influenciada por múltiplos fatores que ocorrem não somente no indivíduo em si, mas entre as populações. A heterogeneidade genética entre os grupos populacionais influencia diretamente os resultados nos estudos de associação genética (MOONESINGHE et al., 2012), sendo já demonstrado que indivíduos de populações mais estudadas e homogêneas, como asiáticos e europeus, e indivíduos de populações pouco estudadas e mais heterogêneas, como americanos, respondem de forma diferente a tratamentos e a susceptibilidade de doenças, entre outros (MARTIN et al., 2017; RAMAMOORTHY et al., 2022).

1.3. ANCESTRALIDADE GENÔMICA E COVID-19

Ademais, a população brasileira é altamente miscigenada e sofre influência genética de diversos grupos populacionais, mas principalmente de europeus, africanos e povos originários, resultando em uma composição genética extremamente heterogênea e um desafio para estudos genéticos de associação, uma vez que a população se apresenta estruturada em relação a contribuição dos diferentes grupos étnicos que se estabeleceram nas distintas regiões geográficas do país (RODRIGUES-SOARES et al., 2018; RODRIGUES et al., 2019; SECOLIN et al., 2019).

Sendo assim, considerando a influência que a ancestralidade genômica populacional possui no desenvolvimento e/ou agravamento de doenças, a sua avaliação genômica em estudos de pesquisa também se faz necessário. Uma das estratégias de realizar esse processo é por meio do uso de Marcadores Informativos de Ancestralidade (MIA) (ENOCH et al., 2006, SANTOS et al., 2010), isto é, por um conjunto de marcadores que são mais frequentes em grupos populacionais específicos, do que em outros o que permite, por cálculo de probabilidade simples, inferir a origem do indivíduo (SHRIVER et al., 2003). Essa estratégia além de permitir a comparação de ancestralidade entre grupos casos e controle, controla o efeito desta em populações miscigenadas.

Nosso grupo de pesquisa publicou em 2010 o primeiro conjunto de MIA do tipo INDEL (Santos et al., 2010), que já foi empregado em estudos de associações em mais de 40 artigos relacionados com doenças (Anexo 1), contribuindo de forma inequívoca para a verificação de associação real entre os marcadores genômicos empregados, assim como, evitando interpretações equivocadas.

Além dessa, é possível avaliar a ancestralidade como um marcador genético por meio da ancestralidade local, definida como a ascendência genética em uma localização genômica de um indivíduo miscigenado (WANG et al., 2020a). Essa abordagem utiliza dados genômicos em larga escala, como o sequenciamento de exomas, que possibilita também a investigação de padrões demográficos associados a doenças (KIMURA & LEMES & NUNES, 2022).

Apesar de poucos estudos associarem a ancestralidade genômica ao maior risco de susceptibilidade ou hospitalização pela COVID-19, Shelton e colaboradores (2021) demonstraram a partir de um estudo GWAS que ancestralidades não-europeias foram associadas a taxas mais elevadas de infecção, sendo que indivíduos com maior contribuição da ancestralidade afro-americana eram 82% mais propensos a serem hospitalizados (SHELTON et al., 2021). Resultados semelhantes também foram observados no estudo de Mathur e colaboradores (2021), onde as ancestralidades asiáticas e africanas possuem maior risco de hospitalização quando comparadas com a ancestralidade europeia (MATHUR et al., 2021).

1.4. BIOINFORMÁTICA

Portanto, para obter os resultados dos estudos supracitados é necessário a geração de uma quantidade massiva de dados das sequências genéticas, e, portanto, uma análise bioinformática para tratar os dados gerados pelo sequenciador é essencial.

A partir da bioinformática é possível tratar esses dados de forma precisa e eficiente por meio do desenvolvimento de pipelines, ou seja, uma cadeia de etapas que orienta e processa progressivamente os dados de sequência por meio de uma série de

transformações utilizando componentes de softwares, bancos de dados e um sistema operacional (LEIPZIG, 2016; ROY et al., 2018).

Sendo assim, o resultado do sequenciamento são diversos fragmentos de DNA por amostra unidos em um arquivo FASTQ (DOLLED-FILHART et al., 2013). Em seguida, esses fragmentos seguem as seguintes etapas para a análise de dados de NGS: chamada de bases, alinhamento dos *reads* ao genoma de referência, e chamada e anotação de variantes (ROY et al., 2018).

A chamada de bases refere-se ao processo de identificação do nucleotídeo presente em cada fragmento, onde cada *read* tem um *score* de qualidade que permite posterior análise. Em seguida, os *reads* são alinhados, ou seja, posicionados corretamente ao longo do genoma de acordo com uma sequência de referência. Como essa etapa fornece a localização dos *reads* no genoma, é possível realizar a proporção dos reads mapeados e verificar a cobertura do sequenciamento, essencial para a próxima etapa (REHM et al., 2013; ROY et al., 2018).

O próximo processo é a chamada de variantes, que consiste em identificar precisamente as diferenças ou variações entre a amostra que está sendo analisada e a sequência do genoma de referência. Nessa etapa, são utilizadas estratégias algorítmicas para reconhecer os diversos os tipos de variantes de sequência, como variantes de nucleotídeo único (SNVs), pequenas inserções e deleções (INDELs), alterações no número de cópias e grandes alterações estruturais (inserções, inversões e translocação) (KOBOLDT, 2020).

Posteriormente, realiza-se a anotação das variantes, onde há uma consulta em diversos bancos de dados a fim de coletar informações sobre cada variante, como a sua localização, se há mudança de aminoácidos, o impacto funcional, se é patogênica e principalmente se há alguma importância clínica (REHM et al., 2013; ROY et al., 2018).

Em seguida, são realizadas análises estatísticas para melhor visualização e compreensão desses dados (disponíveis no formato VCF).

De forma resumida, nosso estudo enfatiza a importância de identificar fatores genéticos, tanto protetivos quanto de risco, do hospedeiro humano na infecção por *SARS-COV-2*. Por conseguinte, a investigação da associação entre fatores genéticos e a gravidade da COVID-19 em uma população amazônica do Norte do Brasil, é essencial para potencializar nossa compreensão a respeito da progressão clínica da doença, a fim de aprimorar os esforços ao sistema de saúde e reduzir a mortalidade dos pacientes infectados pelo SARS-COV-2.

Considerando o exposto, o objetivo deste trabalho é caracterizar marcadores genéticos associados à gravidade clínica da infecção por *SARS-CoV-2* por meio do sequenciamento completo do exoma de pacientes residentes da Região Norte do Brasil, além de avaliar a ancestralidade genômica em relação a gravidade do COVID-19.

2. JUSTIFICATIVA

Os dados epidemiológicos da pandemia da COVID-19 são alarmantes e deixam claro a urgente necessidade de novas estratégias que permitam a melhor compreensão dos mecanismos de ação do vírus e o seu impacto frente às diferentes populações afetadas do mundo, levando ao melhor controle e tratamento da doença (SHELTON et al., 2021).

Tais dados têm demonstrado que a atuação do vírus não é homogênea, ocasionando diferenças significativas nas taxas de transmissão e de morte entre países e continentes. Pacientes infectados por *SARS-CoV-2* podem ter as mais diversas reações, que vão desde a ausência de sintomas clínicos relevantes, até quadros de síndrome respiratória grave, que muitas vezes podem levar o paciente ao óbito, mesmo quando infectados pela mesma cepa.

Portanto, estudos genômicos e epigenômicos relacionados ao hospedeiro podem ser essenciais como uma das estratégias de melhorar compreensão dos mecanismos de ação, de tal modo a esclarecer o impacto das variantes genômicas populacionais sobre a evolução do quadro clínico da doença.

No geral, a análise da variabilidade genética para populações específicas além de ajudar a identificar indivíduos com maior risco de infecção ou de gravidade da doença, também pode auxiliar no desenvolvimento de tratamentos mais eficazes e um maior controle da COVID-19.

Em vista disso, a realização de estudos genômicos, como o exoma, pode favorecer a identificação de novas variantes, evidenciando assim correlações ainda não descritas na literatura, além de contribuírem para o estudo de genes associados a uma condição/doença específica, que por sua vez, cooperam para um melhor entendimento dos fatores que

ocasionam a variabilidade interindividual das respostas a tratamento e progressão da doença.

Considerando que os processos de infecção do *SARS-CoV-2* e do agravamento da COVID-19 são mediados por proteínas do hospedeiro e que o exoma investiga justamente os genes codificadores de proteínas, é evidente a relevância de se utilizar esta técnica para avaliar os possíveis fatores genéticos associados a severidade desta patologia.

Para tal, é necessário o processamento de um grande volume de dados simultâneos gerados pelos estudos de caracterização genômica que necessitam de análises experimentais onerosas e de ferramentas computacionais sofisticadas. Diante deste cenário, fica clara a importância da integração dos dados coletados em repositórios públicos, agregados às informações geradas pelos sequenciamentos de nova geração, com a finalidade de tornar a análise dos dados precisa e compreensível, minimizando erros de interpretação e consequentemente erros de procedimentos.

Portanto, este cenário reforça a importância do presente trabalho que busca investigar a variabilidade presente em todos os genes humanos expressos (investigação de exomas), em populações que habitam a região Amazônica – em especial as do Estado do Pará - de forma a identificar marcadores genéticos específicos que possam estar associados pelo agravamento da infecção, ou por sua proteção na infecção e evolução da doença COVID-19.

3. OBJETIVOS

3.1 OBJETIVO GERAL

Caracterizar o exoma de populações humanas que habitam a região Amazônica em relação a resposta à infecção pelo novo coronavírus *SARS-CoV-2*.

3.2. OBJETIVOS ESPECÍFICOS

I. Identificar os marcadores genômicos, já descritos ou novos, que possam ocasionar maior risco a infecção pelo novo coronavírus *SARS-CoV-2*;

II. Compilar dados públicos de genomas e exomas de populações mundiais e identificar potenciais marcadores genéticos populacionais de risco;

III. Associar marcadores genômicos populacionais observados no presente estudo, em relação aos dois grupos de amostras: pacientes não graves e graves infectados pelo *SARS-CoV-2*;

IV. Correlacionar os marcadores genômicos com os desfechos clínicos dos pacientes com COVID-19;

V. Investigar marcadores genômicos envolvidos em vias metabólicas associadas a COVID-19.

4. METODOLOGIA

O projeto faz parte de dois projetos maiores denominados: (i) "**Vigilância Genômica da COVID-19 baseada em novas estratégias de diagnóstico, prevenção e prognóstico**", coordenado pela Drª. Ândrea Ribeiro-dos-Santos (UFPA) em parceria com a Universidade Federal do Rio Grande do Norte (UFRN) e financiado pela Fundação Amazônia de Amparo a Estudos e Pesquisas (FAPESPA); (ii) "**Rede de pesquisa do genoma do vírus *SARS-CoV-2* na região Amazônica**", coordenado pelo Dr. Guilherme de Oliveira do Instituto Tecnológico Vale (ITV) em parceria com a Universidade Federal do Pará (UFPA), Fundação Oswaldo Cruz (FIOCRUZ) e a Universidade Federal do Rio Grande do Norte (UFRN) e financiado pela Vale.

Ambos os projetos subsidiaram e estão desenvolvendo vários projetos associados com o tema COVID-19, com objetivo de avaliar a expressão de RNAs não codificantes do vírus ("POLIMORFISMOS E EXPRESSÃO DE *circRNAs*: EFEITOS GENÉTICOS E EPIGENÉTICOS NA COVID-19"), a caracterização do transcriptoma tanto do *SARS-CoV-2* quanto do hospedeiro ("CARACTERIZAÇÃO E ANÁLISE DO TRANSCRIPTOMA DE *SARS-CoV-2* E HOSPEDEIRO HUMANO") e identificar as linhagens virais presentes nos pacientes infectados ("VIGILÂNCIA GENÔMICA DO VÍRUS *SARS-CoV-2* EM PACIENTES DA REGIÃO NORTE DO BRASIL"). Especialmente, neste último projeto citado, foi desenvolvido por professores participantes do projeto de "Vigilância Genômica da COVID-19 baseada em novas estratégias de diagnóstico, prevenção e prognóstico", o PipeCoV, *pipeline* específico para genotipar as variantes do *SARS-CoV-2*.

4.1. COMITÊ DE ÉTICA

Todos os procedimentos desta pesquisa foram realizados de acordo com a Declaração de Helsinki e do Código de Nuremberg, obedecidas as normas de pesquisas envolvendo seres humanos (Res. CNS 510/2016) do CNS. O desenvolvimento do estudo foi aprovado pelo Comitê de Ética em Pesquisa do Hospital Barros Barreto (CAAE: 50865721.1.0000.0017).

4.2. CARACTERIZAÇÃO DAS AMOSTRAS

O presente estudo foi desenvolvido em 78 amostras de exomas de indivíduos acometidos pela COVID-19 do Brasil, principalmente aqueles de áreas urbanas e rurais da região Norte, com ênfase no estado do Pará. Todos os participantes desta pesquisa foram testados em relação à presença do novo vírus *SARS-CoV-2* por RT-PCR. Em seguida estas amostras foram divididas entre dois grupos distintos:

I. Pacientes com COVID-19 que manifestaram quadros assintomáticos a leves (alCOV) e não precisaram de internação hospitalar;

II. Pacientes com COVID-19 que evoluíram com quadro grave, assim como aqueles que evoluíram ao óbito (gCOV).

Os critérios de seleção das amostras do grupo grave foram de acordo com a Organização Mundial da Saúde que classifica sintomas graves como dificuldade em respirar (dispneia), perda de fala ou mobilidade e dor no peito. Além desses, incluímos

indivíduos que apresentaram saturação de oxigênio abaixo de 95%, indivíduos hospitalizados, tanto internados em leito normal quanto na UTI.

Todas as amostras biológicas (saliva e swab) dos pacientes participantes foram submetidas a triagem clínica, compreendendo pacientes de ambos os sexos, com idades preferencialmente acima de 18 anos.

A Figura 1 descreve o fluxograma dos procedimentos de biologia molecular e de genômica realizados dentro do Laboratório de Genética Humana e Médica da Universidade Federal do Pará, no presente estudo.

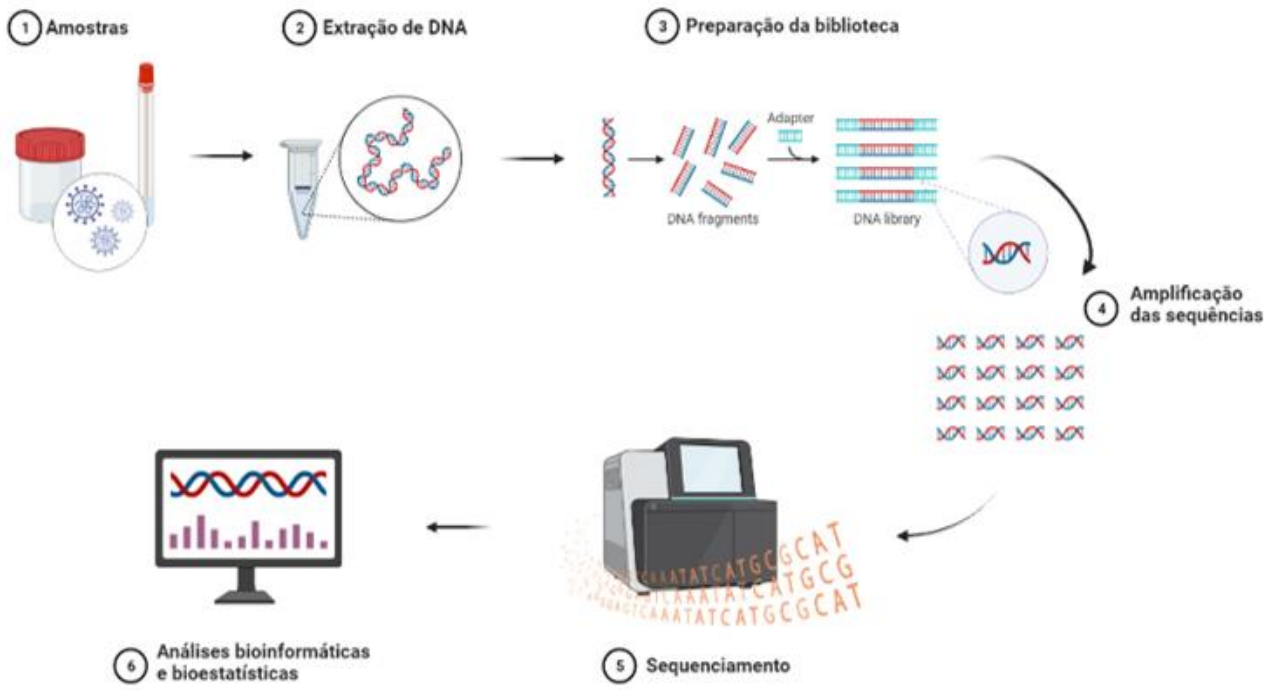


Figura 1. Fluxograma da metodologia empregada no presente projeto.

4.3. ANÁLISE LABORATORIAL

4.3.1 Extração de DNA

Foram selecionadas 78 amostras de saliva e swab de nasofaringe de indivíduos com diagnóstico de COVID-19 confirmado por RT-PCR. As amostras foram separadas em dois grupos: 49 indivíduos com quadro clínico assintomático e/ou leve (alCOV) e 29 indivíduos com quadro clínico grave (gCOV), caracterizado pelos critérios supracitados.

Em seguida, o material genético foi extraído pelos kits de extração *Maxwell® 16 Viral Total Nucleic Acid Purification Kit* (Promega) e *MagMAX DNA Multi-Sample Ultra 2.0 Kit* (Thermo Fisher Scientific). Após a extração, o DNA foi avaliado quanto à integridade e concentração, primeiro realizando eletroforese em gel e em seguida utilizando a metodologia de fluorometria do Qubit 2.0 Fluorometer (Thermo Fisher Scientific), selecionado para dsDNA.

4.3.2 Preparação da biblioteca do exoma

A preparação das bibliotecas foi realizada usando o kit Illumina Exome Panel (Illumina) e Illumina DNA Prep com Exome v2 (Agilent) sendo o input das amostras de 100ng, seguindo o protocolo do fabricante. As bibliotecas foram sequenciadas através de leituras pareadas de 150 pb na plataforma de sequenciamento Illumina NextSeq 550 Systems utilizando o kit NextSeq 500/550 High Output com aproximadamente 12 a 16 amostras cada. A seguir (Figura 2), segue o fluxo da preparação das bibliotecas de acordo com o fabricante.

Figura 2. Workflow da preparação das bibliotecas das amostras investigadas para sequenciamento.

4.4 ANÁLISE BIOINFORMÁTICA

4.4.1. Programas e ferramentas

4.4.1.1. Trimmomatic

Para realizar o controle de qualidade dos dados brutos é utilizada a ferramenta Trimmomatic v.0.36 que performa uma variedade de etapas de trimagem específica para dados gerados pelo sequenciador da Illumina tanto paired-end como single-end. Com essa ferramenta, é possível remover leituras com baixa qualidade (filtering), remover bases com extremidades de baixa qualidade (trimming) ou sequências de adaptadores ou primers (clipping) (BOLGER & LOHSE & USADEL, 2014).

4.4.1.2. BWA MEM

O Burrows-Wheeler Aligner (BWA) é um programa para o a indexação dos genomas e posterior mapeamento de sequências contra um genoma de referência. Atualmente é mais utilizado o método BWA-MEM para alinhamento das leituras geradas por sequenciadores da Illumina, além de ser mais preciso e mais eficiente (LI & DURBIN, 2009; LI, 2013).

4.4.1.3. SAMtools

O SAMtools é um pacote de software utilizado para analisar e manipular alinhamentos no formato SAM/BAM. Com essa ferramenta é possível classificar e mesclar alinhamentos, mostrar alinhamentos com melhor visualização em texto e como

possui um algoritmo Bayesiano, é praticável realizar chamada de variantes curtas (LI et al., 2009).

4.4.1.4. PICARD Tools

O PICARD Tools é uma coleção de ferramentas de linhas de comando baseadas em JavaScript, desenvolvida pelo *Broad Institute*, para manipular dados de sequenciamento de alto rendimento em arquivos SAM/BAM. Essa ferramenta permite indexar os arquivos e remover duplicatas de PCR, além de fazer a conversão do arquivo de SAM para BAM, para ser utilizado na próxima etapa (broadinstitute.github.io/picard/).

4.4.1.5. GATK

O *Genome Analisys ToolKit* (GATK) é um *framework* desenvolvido em Java pelo *Broad Institute* que possui diversas ferramentas com o objetivo de analisar o genoma com foco na análise de variantes, genotipagem e garantia da qualidade dos dados. Existem diversos *workflows* que são utilizados de acordo com o foco do estudo (VAN DER AUWERA; O'CONNOR, 2020).

4.4.1.6. SnpEff

O SnpEff é uma ferramenta que permite a predição do efeito causado pelas variantes com base na sua localização genômica. Além de realizar a anotação de SNPs e INDELs, ela é capaz de predizer mutações em regiões de splicing, qual o efeito funcional (high, moderate, low ou modifier), entre outras. O arquivo de saída é no formato VCF e contém as mutações anotadas (CINGOLANI et al., 2012).

4.4.2. Bancos de referências

4.4.2.1. ClinVar

O ClinVar foi desenvolvido com o intuito de associar variantes genéticas a fenótipos clínicos, sendo possível observar a significância clínica, o tipo de alelo e a qual doença aquele polimorfismo está associado (LANDRUM et al., 2018). Para o presente trabalho, esse banco de dados foi coletado do NCBI (National Cancer for Biotechnology Information), disponível em: Index of /pub/clinvar (nih.gov).

4.4.2.2. COSMIC

O COSMIC (Catalogue Of Somatic Mutations In Cancer) é o banco de dados mais abrangente e íntegro no quesito impacto das mutações somáticas do câncer em humanos, uma vez que combina dados de painéis genéticos específicos e estudos genômicos em larga escala (TATE et al., 2019). Esse banco está disponível em: Download Files (sanger.ac.uk).

4.4.2.3. dbSNP

O intuito do dbSNP, desenvolvido pelo NCBI, é facilitar os estudos de associação em larga escala, mapeamento de genes, de farmacogenômica e biologia evolutiva por meio de um catálogo geral de variação do genoma de diversas espécies (SHERRY, 2001). É possível realizar o download deste banco em: Home - SNP - NCBI (nih.gov).

4.4.2.4. BRCA1

O banco BRCA1 disponibiliza variantes de nucleotídeo único (SNVs) em 13 éxons que codificam domínios funcionalmente críticos do gene *BRCA1*, utilizado neste trabalho a fim de identificar variantes germinativas com perda de função (FINDLAY et al., 2018).

4.4.2.5. ExAC

O Exome Aggregation Consortium (ExAC) é um consórcio de pesquisadores cujo objetivo é agregar dados de sequenciamento de exoma de diversos projetos em larga escala a fim de disponibilizar tais dados de forma mais acessível para a comunidade cientifica (KARCZEWSKI et al., 2020). Atualmente está disponível no gnomAD (Genome Aggregation Database) em: Downloads | gnomAD (broadinstitute.org).

4.4.3. Processamento de reads e chamada de variantes

A análise dos dados de sequenciamento recebidos foi realizada utilizando o *pipeline* desenvolvido por professores da UFRN, descrito nas linguagens *ShellScript* e *R* que utiliza softwares feitos em JavaScript como o *GATK*, *PICARD* e *TRIMMOMATIC*. Este pipeline é dividido em quatro partes: Controle de Qualidade, Alinhamento, Pré-processamento, Chamada de Variantes e Anotação das variantes.

A partir do *base-calling* realizado pelo sequenciador, é analisado a qualidade e a cobertura dos reads gerados pela ferramenta *FASTQC* e em seguida as sequências de baixa qualidade são eliminadas pelo *Trimmomatic v.0.36* (BOLGER et al., 2014).

Em seguida, as sequências foram alinhadas ao genoma de referência (GRCh38) utilizando a ferramenta *BWA MEM* v.0.7 (LI & DURBIN, 2009; LI, 2013). O arquivo foi indexado e ordenado pela ferramenta *SAMtools v.1.7* (LI et al., 2009) e em seguida, pré-processado pela ferramenta *Picard v.2.27.3* (broadinstitute.github.io/picard/), onde foi realizado a classificação das reads por coordenadas do genoma de referência seguida da sua reordenação e a remoção de duplicatas.

A chamada de variantes foi realizada de acordo com o *GATK Best Pratice Guidelines v. 4.1.9.0* (VAN DER AUWERA; O'CONNOR, 2020) onde as amostras passaram pelos seguintes processos: recalibração do índice de qualidade base (BaseRecalibrator e ApplyBQSR); chamada de SNPs e INDELs por remontagem local de haplótipos (HaplotypeCaller); importação dos arquivos em .VCF para o banco GenomicsDB(GenomicsDBImport); genotipagem conjunta, que permite a identificação das variantes dentro das regiões-algo e sua junção em uma coorte (GenotypeGVCFs); recalibração de bases (SNPs e INDELs) para pontuar a qualidade da variante para fins de filtragem (VariantRecalibrator -mode SNP -mode INDEL); e filtragem de variantes (ApplyVQSR -mode SNP -mode INDEL).

Os resultados obtidos foram anotados usando *SnpEff v.5.1, COSMIC v.96, dbnsfp v4.2* e *BRCA1* para incluir o impacto e predição funcional das variantes, a importância clínica de acordo com o banco de dados *ClinVar* (LANDRUM et al., 2018) e anotações de frequência alélica do *ExAC* e *gnomAD* (EXOME AGGREGATION CONSORTIUM et al., 2016; KARCZEWSKI et al., 2020). Segue abaixo, Figura 3, o fluxograma resumido do *pipeline* utilizado para as análises de bioinformática do exoma.

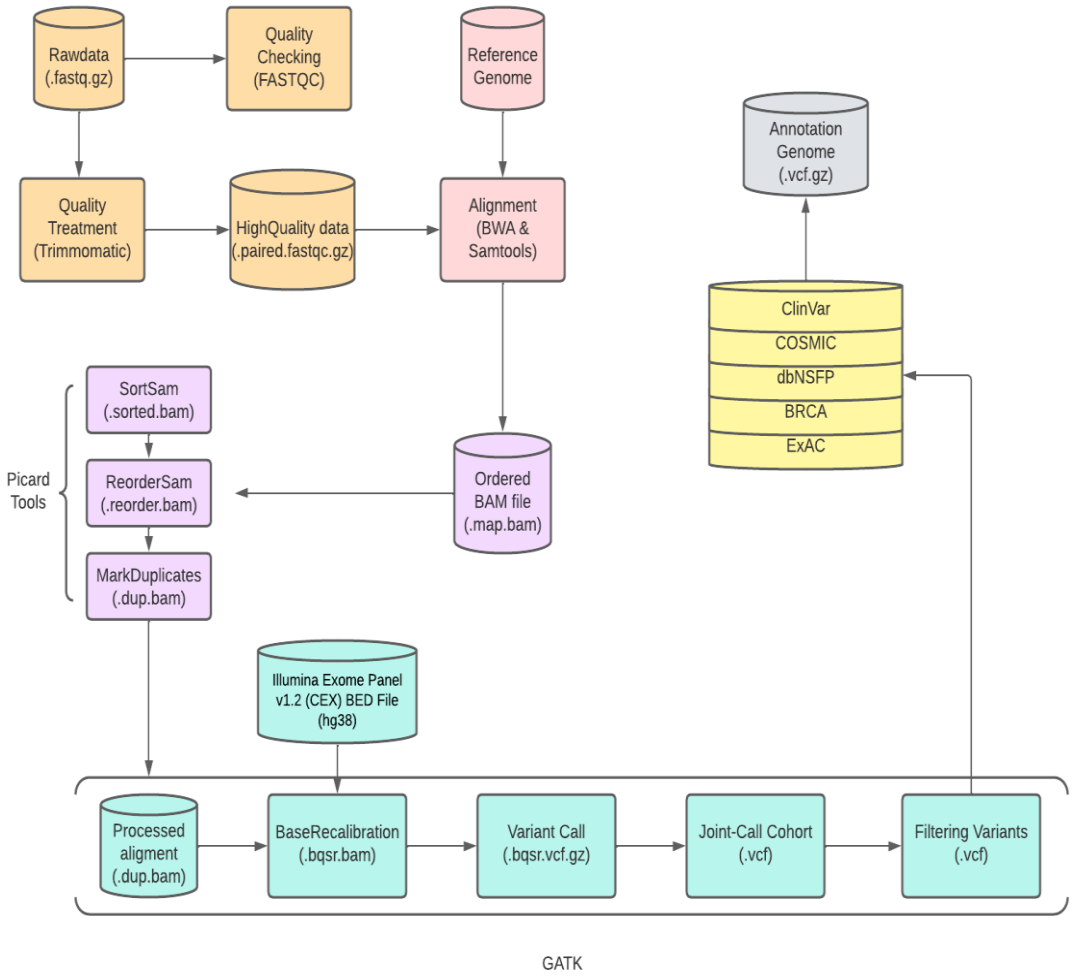


Figura 3. Pipeline do Exoma. O fluxograma exibe todos os principais passos para análise do exoma. Em laranja, controle de qualidade dos *reads*. Em rosa, mapeamento dos *reads* utilizando genoma de referência (Hg38). Em roxo, pré-processamento das amostras utilizando a ferramenta *Picard*. Em azul, chamada de variantes por amostras utilizando o pipeline da ferramenta *GATK*. E em amarelo, os bancos de dados utilizados para anotação e identificação das variantes.

4.5. ANÁLISES GENÔMICO-POPULACIONAIS

Os dados deste projeto foram reunidos aos dados dos projetos *1,000 Genomes* e *HGDP*, totalizando 4.104 indivíduos. O controle de qualidade dos dados genômicos foi realizado por meio dos softwares *BCFTools v1.8, VCFTools v0.1.16* e *PLINK v1.90b4.1*. Foram excluídas variantes com taxa de *missing data* superior à 0,001 e indivíduos com taxas de dados faltantes superiores à 0,1.

Os padrões de estruturação populacional foram avaliados a partir dos *softwares Eigenstrat v8.0.0* e *Admixture v1.3*.0. Para ambas, utilizou-se um subconjunto das variantes obtidas a partir da exclusão de variantes em alto desequilíbrio de ligação (pruning). Na análise de PCA, foram utilizadas todas as populações dos conjuntos de dados *1,000 Genomes, HGDP* e pacientes com COVID. Para as análises do Admixture, recorremos à duas abordagens: i) análises supervisionadas admitindo um modelo tri-híbrido e as populações Ibérica (IBS - 1KG), Iorubá (YRI/AFR - 1KG) e Ameríndios (AMR - HGDP); ii) análises não supervisionadas utilizando as mesmas populações do modelo supervisionado e outra rodada englobando todas as populações dos painéis do *1,000 Genomas*, do *HGDP* e dos pacientes com COVID, ambas variando o número de clusters (K) de 2 a 12.

O ambiente *R v4.2.2* foi utilizado para as análises de epidemiologia genética e controle de qualidade dos dados genéticos. Uma rodada adicional de controle de qualidade dos dados de sequenciamento e chamada de variantes foi realizada com os pacotes *vcfR v1.13.0, variantAnnotation 1.42.1* e *seqVarTools 1.34.0*. Os dados de VCF dos pacientes de COVID foram convertidos ao formato GDS (Genetic Data Structure) e

este utilizado nas análises *downstream*. A análise de identidade por descendência (IBD - Identity-by-descent) e o cálculo da matriz de relacionamento genético foram obtidos a partir do método KING implementado no pacote *SNP Relate v1.30.1*. A partir da matriz de relacionamento genético (GRM), foi realizada a análise de componentes principais (PCA) corrigida para os pacientes de COVID. O pacote *GENESIS v2.26.0* foi utilizado para o cálculo dos coeficientes de *kinship*, coeficientes de *inbreeding* e das probabilidades de compartilhamento de IBD. A partir das análises de parentesco foi excluído um indivíduo de um par com valores de *kinship* superiores à 0,004. Com o intuito de predizer as covariáveis associadas à gravidade da doença, foi realizado um modelo linear generalizado (GLM) envolvendo sexo, idade, ancestralidade ameríndia, africana e europeia. O pacote GENESIS foi ainda utilizado para os testes individuais de associação genética, sendo incluídas as covariáveis idade e os cinco primeiros componentes principais, assim como para o teste de associação genética agregado realizando a agregação dos dados por genes.

4.6. ANÁLISES ESTATÍSTICAS

Todas as análises estatísticas e gráficos foram performados no programa *R v.4.2.2*. As variáveis clínicas foram analisadas por Teste T (idade), Qui-quadrado com correção de Monte Carlo (sexo), teste exato de Fisher para variáveis categóricas e o teste Shapiro-Wilk como teste de normalidade. Para a análise descritiva das variantes observadas no exoma, foram realizadas análises bivariadas. Foi realizada análise de regressão logística para avaliar a associação entre os polimorfismos e a gravidade clínica da COVID-19, sendo considerados fatores de confusão: idade, linhagem viral e comorbidade. O p-valor foi considerado estatisticamente significativo quando igual ou abaixo de 0.05.

5. RESULTADOS

5.1. ANÁLISE CLÍNICA

Os dados demográficos e clínicos dos pacientes estão presentes na Tabela 1. No total, foram incluídos 78 indivíduos no estudo, sendo 29 do grupo gCOV (quadro grave e/ou evoluiu ao óbito) e 49 do grupo alCOV (quadros assintomáticos e/ou leves). Todos os indivíduos foram positivos para COVID-19, na época do diagnóstico, confirmados por RT-qPCR. Em relação ao valor do CT observado na RT-qPCR, não houve diferença nas médias observadas em cada grupo (gCOV = 21.3; alCOV=27.03, pvalor = 0.44).

A média de idade do grupo gCOV foi de 57.2 (± 11.5) e a do grupo alCOV de 42.1 (± 15.1), sendo observada associação significativa entre a idade e a gravidade clínica da COVID-19 (*pvalor* = 0.005; IC 95% = 8.98 – 21.2). Em relação ao sexo, houve prevalência do sexo feminino no grupo grave (58,6%), entretanto não foi observado diferença significativa (*pvalor* = 0.48).

Dentre as treze linhagens identificadas no estudo, duas foram mais prevalentes: B.1.1 (*pvalor* = 0.42; OR= 1.6; IC 95% = 0.52 – 5.23) e P.1 (*pvalor* = 0.47; OR= 0.6; IC 95% = 0.23 – 1.89), porém nenhuma foi associada significativamente com a gravidade clínica da COVID-19.

Tabela 1. Dados bioantropológicos e clínicos dos pacientes com COVID-19.

Variáveis	gCOV (n=29)	alCOV (n =49)	OR (95% IC)	p-valor
Idade	57.2 (± 11.5)	42.1 (± 15.1)	(8.98 - 21.2)	**0.005***
Sexo				0.48
Feminino	17 (58,6%)	24 (48,9%)		
Masculino	12 (41,4%)	25 (51,02%)		
Linhagem				**0.05***
B.1.1	10	11	1.6 (0.52 – 5.23)	0.42
P.1	11	22	0.6 (0.23 – 1.89)	0.47
Comorbidades	22	12	9.3 (2.97 – 33.05)	**0.01***
HAS	14	4	10.1 (2.66 – 48.9)	**<0.01***
Obesidade	4	1	7.4 (0.6 – 385.02)	0.06
Doença Cardíaca Crônica	4	3	2.4 (0.37 – 17.8)	0.41
Sobrepeso	9			0.05
Sintomas	29	33		**0.002***
Febre	3	11	2.47 (0.31 – 17.2)	0.35
Tosse	5	20	3.42 (0.49 – 39.3)	0.22
Dor de Garganta	2	21	0.52 (0.04 – 3. 57)	0.65
Comprometimento Pulmonar				
25-50%	8			
>50%	10			
>75%	3			
CT qPCR	21.38	27.03		0.44

gCOV: pacientes com sintomas graves. alCOV: pacientes assintomáticos e com sintomas leves. *Teste T. ** Teste Qui-Quadrado. ***Teste Exato de Fisher.

Em relação às comorbidades, foi encontrada uma associação significativa e com o risco de nove vezes maior de desenvolver a forma grave (*pvalor* = 0.01; OR= 9.3; IC 95% = 2.97 – 33.05) caso o indivíduo apresentasse alguma doença prévia. Das comorbidades analisadas, apenas a HAS (*pvalor* <0.01; OR= 10.1; IC 95% = 2.66 – 48.9) apresentou significância estatística entre os dois grupos, apesar da obesidade apresentar valores elevados (*pvalor* =0.06; OR= 7.4; IC 95% = 0.6 – 385.02).

Dentre as manifestações clínicas, as mais prevalentes entre o grupo que apresentou sintoma não grave foram: dor de garganta (43,7%), tosse (41,6%) e febre (22,9%). Visto que a maioria dos pacientes do grupo grave se apresentarem hospitalizados dentro da Unidade de Tratamento Intensivo - UTI (65,5%), não foi possível coletar informações completas sobre os sintomas. Entretanto, naqueles com os sintomas descritos, todos apresentaram dispneia, a maioria (75%) estava com a saturação de oxigênio abaixo de 95%, com alteração na tomografia (79,3%) e com comprometimento pulmonar na faixa de >50% como mais predominante (43,4%).

5.2 ANÁLISE DESCRITIVA DAS VARIANTES

Nos exomas dos pacientes investigados foram observadas 89.994 variantes no total, das quais se destacam 1.461 (1,62%) variantes ainda não descritas na literatura mundial. Das variantes classificadas (Figura 4), ao todo foi observado 1.811 (2,01%) patogênicas/levemente patogênicas (de acordo com o banco de dados *ClinVar*) ou danosas (de acordo com o algoritmo de predição *MetaSVM*), das quais foram observadas em relação ao i) tipo de efeito: 7 substituições *frameshift*, 9 mutações *stop gained*, 7 intrônicas, 1704 não sinônimas na região exon e 76 na região splicing; ii) impacto

funcional: 18 variantes de alto impacto, 10 de impacto modificador e 1780 de impacto moderado; iii) classe funcional: 1780 *missense*, 9 *nonsense* e duas silenciosas (Figura 4).

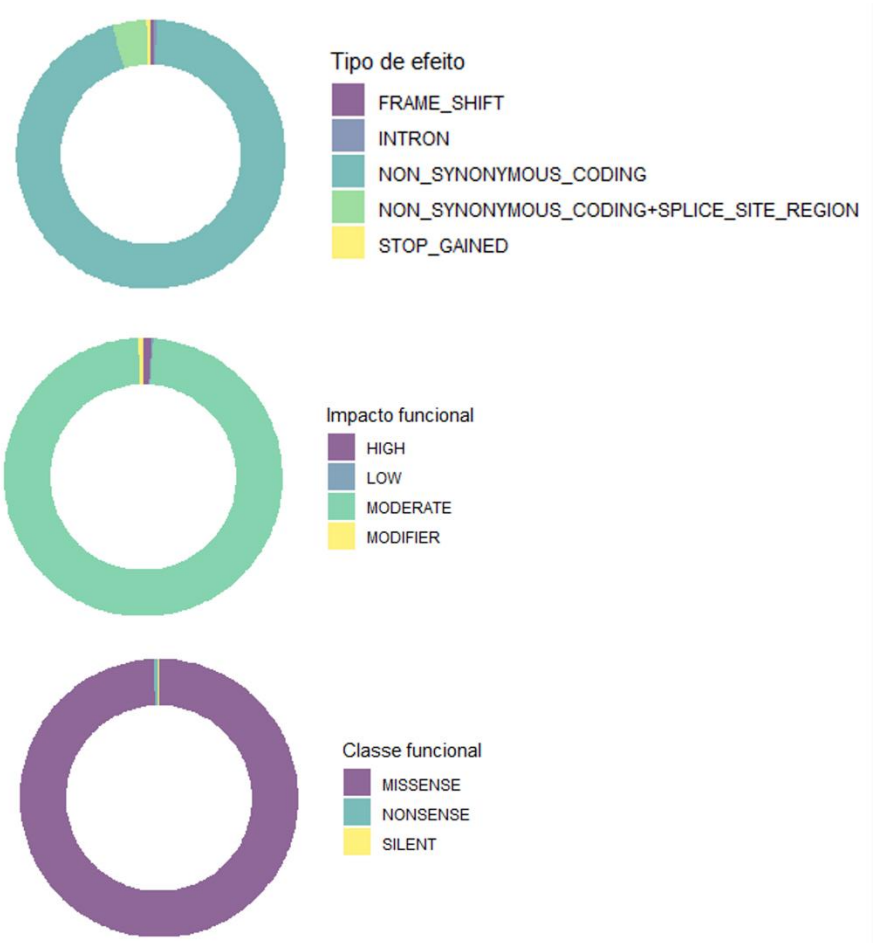


Figura 4. Distribuição das variantes genômicas humanas observadas nos exomas dos pacientes com COVID-19.

Em relação as variantes classificadas como patogênicas e/ou danosas, 355 foram observadas somente no grupo grave, 584 apenas no grupo não grave e 111 variantes comuns aos dois grupos (Figura 5).

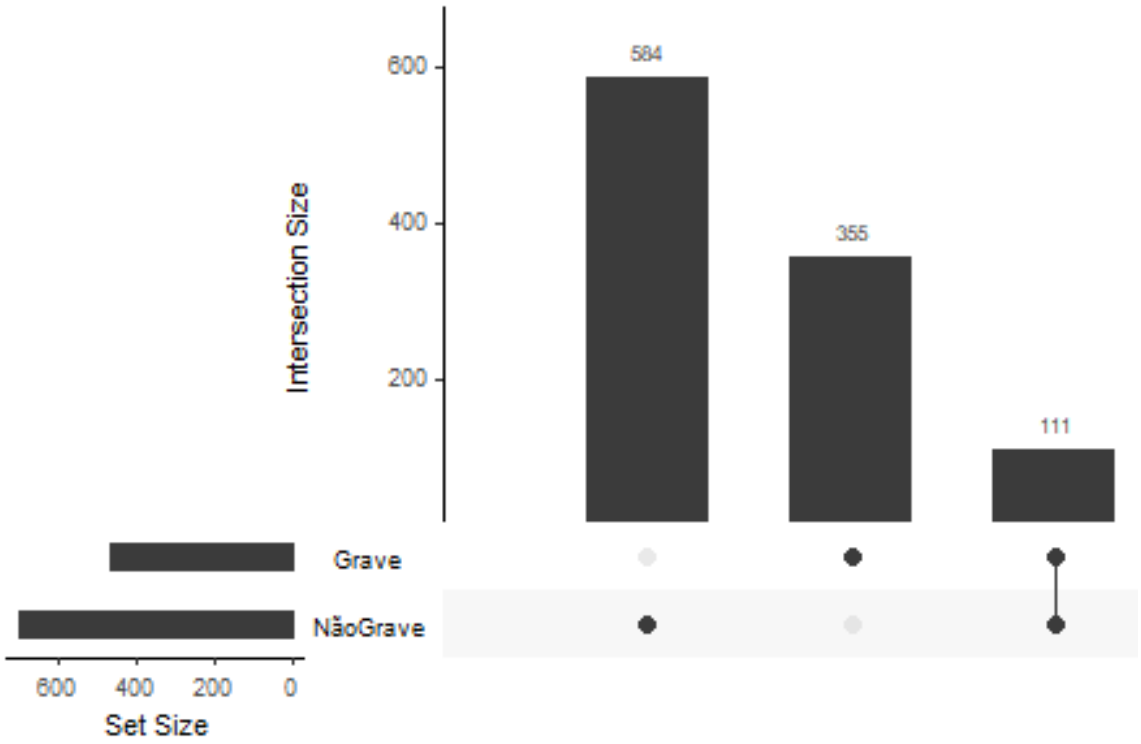


Figura 5. Distribuição das variantes genômicas humanas observadas nos exomas dos pacientes com COVID-19, divididas entre os dois grupos investigados (graves e não graves).

No grupo grave, foram observadas 709 variantes patogênicas em relação ao tipo de efeito das mutações: 657 variantes não sinônimas na região do éxon, 41 variantes na região de *splicing*, 2 substituições *frameshift* e 2 mutações *stop gained*. Em relação ao impacto funcional, foram encontradas: 698 de impacto moderado, 5 de alto impacto, 3 de baixo impacto e 3 de impacto modificador. Em relação a classe funcional, foram identificadas: 698 missense, 2 nonsense e 2 silenciosas. É possível observar a proporção dessas variantes na Figura 6A, B, e C.

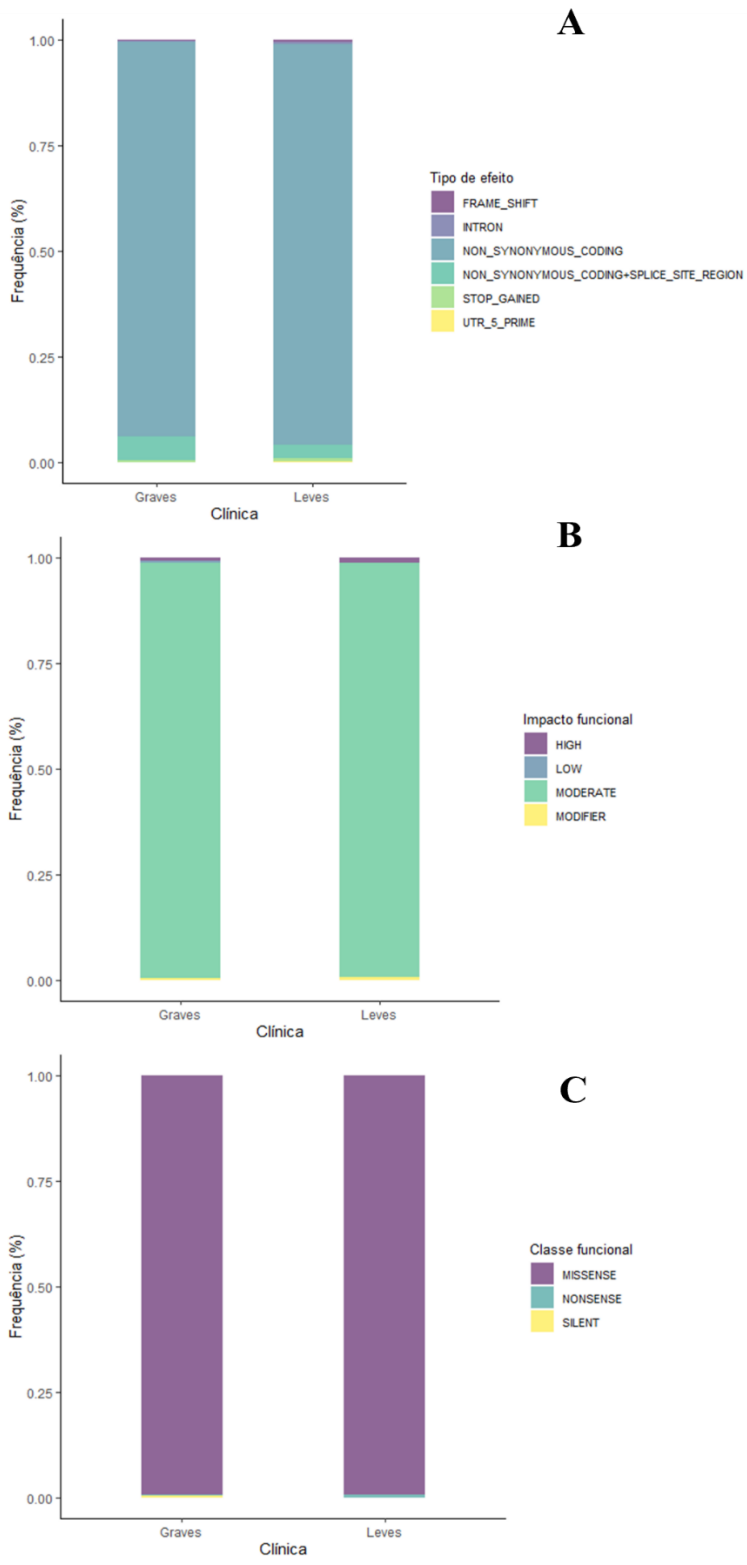


Figura 6. Proporção dos tipos de variantes observadas, em relação ao tipo de efeito, impacto funcional e classe funcional, nos dois grupos de pacientes que apresentaram diagnóstico positivo para COVID-19.

No grupo de pacientes leves, em relação ao tipo de efeito das 1.102 variantes patogênicas observadas: 1.047 são variantes não sinônimas na região do éxon, 35 variantes na região de *splicing*, 7 são mutações *stop gained* e 5 substituições *frameshift*. Em relação ao impacto funcional, foram encontradas: 1.082 de impacto moderado, 13 de alto impacto e 7 de impacto modificador. Em relação a classe funcional foram identificadas 1.082 variantes *missenses* e 7 variantes *nonsens*e, como é possível observar na Figura 6.

No que se refere às dezoito mutações de alto impacto encontradas, ainda não há evidências científicas robustas que possam correlacionar as variantes em si com algum fenótipo clínico (Tabela 2). As informações sobre as sete substituições *frameshift* também se encontram na Tabela 2.

Todavia, ao se avaliar os genes e as vias metabólicas que essas mutações participam, é possível observar que os pacientes que formam o grupo grave, apresentam um conjunto de mutações em genes associados as vias da beta-oxidação mitocondrial de ácidos-graxos (*HADHA*), adesão celular (*MPZL2*) e do reparo celular por recombinação homóloga (*PALB2*) que podem estar associadas a mecanismos de patogênese do *SARS-CoV-2* ainda não descritos.

Por outro lado, aqueles pacientes que se enquadram no grupo não grave, foram observadas mutações em genes que participam dos processos metabólicos envolvidos com a resposta imune inata (*SAMD9* e *ADA2*) e com o reparo celular pela via de recombinação homóloga (*NSMCE2*), que está associada com potencial de defesa contra a resposta viral.

Tabela 2. Mutações de alto impacto observadas nos grupos grave e não grave.

	CHROM	POS	ID	REF	ALT	GENE	EFEITO	classe	AA	TIPO	PROCESSOS METABÓLICOS ASSOCIADAS AO GENE
GRAVES											
	chr2	26192390	rs796051971	C	CTGAT	HADHA	frame_shift		p.Glu641fs/c.1916_1919dupATC	INDEL	Beta-oxidação mitocondrial de ácidos-graxos; Resposta a insulina
	chr7	2586272	rs773701437	TTGTCC	T	IQCE	frame_shift		p.Val301fs/c.895_904del	INDEL	Regulação ciliar na via de sinalização Hedgehog
	chr10	84201806	.	G	A	CDHR1	splice_site_acceptor		c.526-1G>A	SNV	Adesão celular
	chr11	118260094	rs200462584	G	A	MPZL2	stop_gained	nonsense	p.Arg182*/c.544C>T	SNV	Adesão celular
	chr16	23636043	.	G	C	PALB2	stop_gained	nonsense	p.Ser168*/c.503C>G	SNV	Processo apoptótico; Reparo celular via recombinação homóloga
NÃO GRAVES											
	chr1	150799607	rs74315303	G	A	CTSK	stop_gained	nonsense	p.Arg241*/c.721C>T	SNV	Autofagia mitocondrial; Quebra da matriz extracelular
	chr2	232540049	rs797044677	A	AC	CHRNG	frame_shift		p.Asn40fs/c.117dupC	INDEL	Transmissão sináptica
	chr3	189974638	rs895635914	G	A	P3H2	stop_gained	nonsense	p.Gln458*/c.1372C>T	SNV	Processo metabólico do colágeno; Regulação negativa de proliferação celular
	chr3	44901672	rs139860990	G	A	TGM4	stop_gained	nonsense	p.Trp269*/c.806G>A	SNV	Catalização da ligação proteica
	chr4	158680553	rs773668457	C	T	ETFDH	stop_gained	nonsense	p.Arg41*/c.121C>T	SNV	Cadeia de transporte de elétrons; Resposta ao estresse oxidativo
	chr5	178991553	rs752625077	AC	A	GRM6	frame_shift		p.Val243fs/c.727delG	INDEL	Transmissão sináptica
	chr6	63999114	rs758109813	TG	T	EYS	frame_shift		p.Pro2265fs/c.6794delC	INDEL	Detecção de estímulo envolvido na percepção visual; Regeneração muscular
	chr7	93102535	.	CGCCTT	C	SAMD9	frame_shift		p.Arg1187fs/c.3558_3562delAA	INDEL	Fusão da vesícula endossomal; Resposta imune inata
	chr8	132882609	rs121912648	C	T	TG	stop_gained	nonsense	p.Arg296*/c.886C>T	SNV	Formação do hormônio da tireoide
	chr8	125182182	rs757613817	AT	A	NSMCE2	frame_shift		p.Ser116fs/c.346delT	INDEL	Reparo celular via recombinação homóloga; Regulação telomérica
	chr11	78106895	rs376161880	G	A	ALG8	stop_gained	nonsense	p.Arg364*/c.1090C>T	SNV	Glicosilação
	chr12	102855316	rs199475575	G	A	PAH	stop_gained	nonsense	p.Arg176*/c.526C>T	SNV	Processo biosintético de aminoácidos
	chr22	17188449	rs139750129	T	C	ADA2	splice_site_acceptor		c.973-2A>G	SNV	Resposta imune inata; Processo catabólico da adenosina

Em relação às 1780 mutações *missense* observadas, foi possível visualizar a localização genômica de apenas 72 variantes, sendo 40 encontradas em regiões intrônicas, 12 em regiões exônicas, 7 em regiões intragênicas e 5 em regiões 3'UTR e 2 em regiões 5'UTR do RNA mensageiro (Figura 7).

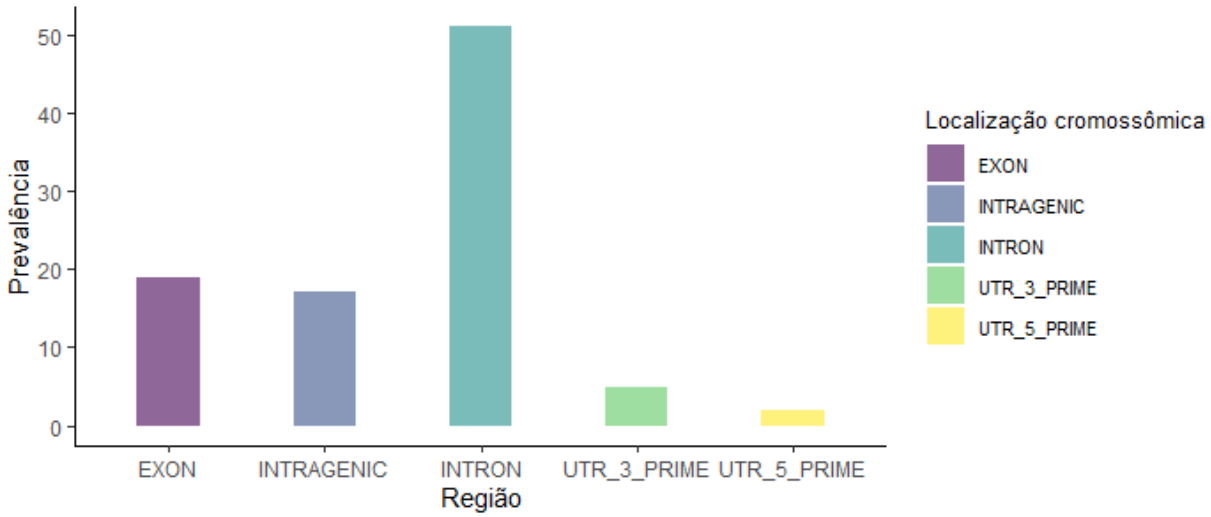


Figura 7. Distribuição das variantes missenses encontradas de acordo com a localização genômica.

No que se refere às mutações *missenses* presentes em regiões exônicas, o foco desse estudo, é possível observar que todas são não sinônimas de impacto moderado (Tabela 3). A variante rs1799945 está em destaque pois apareceu oito vezes no grupo não grave e uma vez no grupo grave. A maioria das variantes observadas ainda não foram descritas na literatura e por isso foi avaliado os genes e as vias metabólicas que essas mutações estão envolvidas.

A partir disso, é possível observar que os pacientes que formam o grupo grave, apresentam um conjunto de mutações em genes associados as vias do processo catabólico

da selenocisteína (*SCLY*), da regulação negativa de citocinas da célula T (*HFE*), da ativação (*HOXA9*) e da regulação da proliferação (*FUT2*) de células endoteliais. No que se refere ao grupo não grave, observa-se variantes envolvidas nos processos de beta-oxidação de ácidos graxos (*ACAD11*), de regulação negativa da entrada viral na célula (*TRIM59*), de regulação negativa de citocinas da célula T (*HFE*), de desenvolvimento de órgãos (PTCH1), de queratinização (*KRT7*), de fertilização (*CATSPER2*) e de desenvolvimento do sistema neurológico (*P2RX5*) (Tabela 3).

Tabela 3. Mutações *missenses* presentes em regiões exônicas observadas nos grupos grave e não grave.

	CHROM	POS	ID	REF	ALT	GENE	EFEITO	IMPACTO	AA	PROCESSOS METABÓLICOS ASSOCIADOS AO GENE
GRAVES										
	chr2	238064441	rs776158924	G	T	SCLY	NON_SYNONYMOUS_CODING	MODERATE	p.Trp58Cys/c.174G>T	Processo catabólico da selenocisteína
	chr6	26090951	rs1799945	C	G	HFE	NON_SYNONYMOUS_CODING	MODERATE	p.His63Asp/c.187C>G	Regulação negativa de citocinas da célula T; Homeostase celular do íon de ferro
	chr7	27164945	rs138291525	G	T	HOXA9	NON_SYNONYMOUS_CODING	MODERATE	p.Ser171Arg/c.513C>A	Ativação da célula endotelial; Desenvolvimento de glândulas sexuais
	chr19	48703368	rs1800022	C	T	FUT2	NON_SYNONYMOUS_CODING	MODERATE	p.Arg138Cys/c.412C>T	Regulação da proliferação de células endoteliais; Glicolisação proteica
	chr19	48703291	rs200157007	C	T	FUT2	NON_SYNONYMOUS_CODING	MODERATE	p.Pro112Leu/c.335C>T	Regulação da proliferação de células endoteliais; Glicolisação proteica
NÃO GRAVES										
	chr3	132559063	rs138611343	G	A	ACAD11	NON_SYNONYMOUS_CODING	MODERATE	p.Arg751Cys/c.2251C>T	Beta-oxidação de ácidos graxos
	chr3	160438920	rs78209671	T	G	TRIM59	NON_SYNONYMOUS_CODING	MODERATE	p.Gln88His/c.264A>C	Resposta imune inata; Regulação negativa da entrada viral na célula
	chr6	26090951	rs1799945	C	G	HFE	NON_SYNONYMOUS_CODING	MODERATE	p.His63Asp/c.187C>G	Regulação negativa de citocinas da célula T; Homeostase celular do íon de ferro
	chr6	26090957	rs1800730	A	T	HFE	NON_SYNONYMOUS_CODING	MODERATE	p.Ser65Cys/c.193A>T	Regulação negativa de citocinas da célula T; Homeostase celular do íon de ferro
	chr9	95467191	rs201125580	C	T	PTCH1	NON_SYNONYMOUS_CODING	MODERATE	p.Val829Met/c.2485G>A	Desenvolvimento de órgãos; Regulação do ciclo mitótico celular
	chr12	52245536	rs141666299	G	A	KRT7	NON_SYNONYMOUS_CODING	MODERATE	p.Arg370Gln/c.1109G>A	Queratinização; Organização do citoesqueleto
	chr15	43632281	rs763876192	T	A	CATSPER2	NON_SYNONYMOUS_CODING	MODERATE	p.Arg497Ser/c.1491A>T	Fertilização
	chr17	3688639	rs151100959	C	T	P2RX5	NON_SYNONYMOUS_CODING	MODERATE	p.Gly292Arg/c.874G>A	Transdução de sinal; Desenvolvimento do sistema neurológico

5.3 ANÁLISE DA ANCESTRALIDADE GENÔMICA

A partir de dados genômicos obtidos por meio do sequenciamento do exoma, foi avaliado os padrões de estruturação populacional presentes na amostra de pacientes de COVID-19. Para as análises do *software* Admixture, foi realizado a análise supervisionada em um modelo tri-híbrido considerando amostras de indivíduos dos grupos: europeus representados pelas populações da península Ibérica (IBS - 1KG), africanos representados pelo povo Iorubá (YRI/AFR - 1KG) e de nativos americanos representados pelos povos originais da Amazônia (AMR - HGDP) (Figura 8).

Na amostra de pacientes com COVID-19, quanto à ancestralidade individual, é possível observar uma maior contribuição da ancestralidade de povos originais (0,62 – cor vermelha), seguida da ancestralidade europeia (0,30 – cor azul) e da ancestralidade africana (0,08 – cor verde). Ressalta-se que 18 indivíduos apresentam ancestralidade ameríndia superior a 90%, enquanto apenas um único indivíduo apresentou ancestralidade europeia maior que 90% e nenhum apresentou ancestralidade majoritariamente africana.

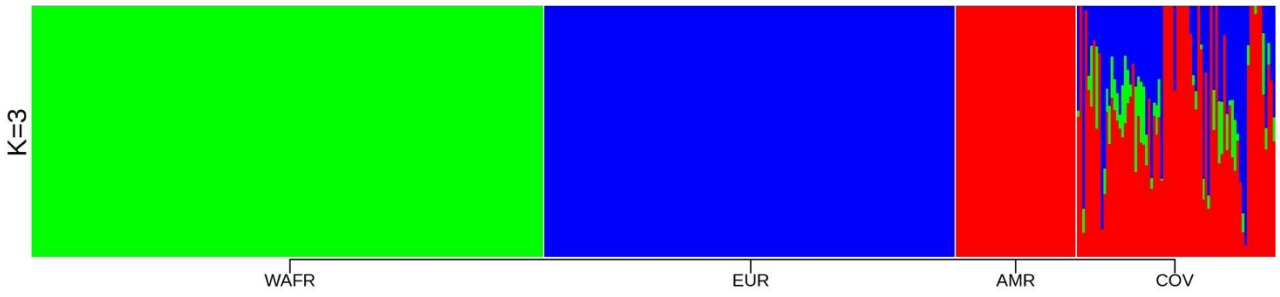


Figura 8. Análise supervisionada dos padrões de estruturação populacional presentes na amostra do estudo. WAFR - África Ocidental (YRI), EUR - Europa (IBS), AMR – Povos Originais, e COV - Pacientes acometidos pela COVID-19.

A partir da Análise de Componentes Principais (PCA) foi possível verificar os padrões de diversidade genética e estruturação populacional dos pacientes de COVID-19 em relação à populações de diferentes regiões biogeográficas. Além disso, os cinco primeiros componentes principais (PCs) foram utilizados para a correção da estrutura populacional nos estudos de associação. Para tal, foram utilizadas todas as populações dos conjuntos de dados *1,000 Genomes*, *HGDP* e os pacientes com COVID.

Como é possível observar na Figura 9, os pacientes acometidos pela COVID-19 (em laranja) apresentam um padrão de diversidade genética similar aos dos indivíduos Latino-Americanos depositados no banco de dados do painel do *1,000 Genomes* . Ainda que exista uma grande diversidade na contribuição das populações Nativas Americanas, Européias e Africanas na formação das populações americanas, a amostra de indivíduos brasileiros da região Amazônica, acometidos por COVID-19, apresenta maior contribuição dos povos originais (nativos americanos) e de europeus, com baixa contribuição africana, fenômeno também observado em outras populações da América Latina.

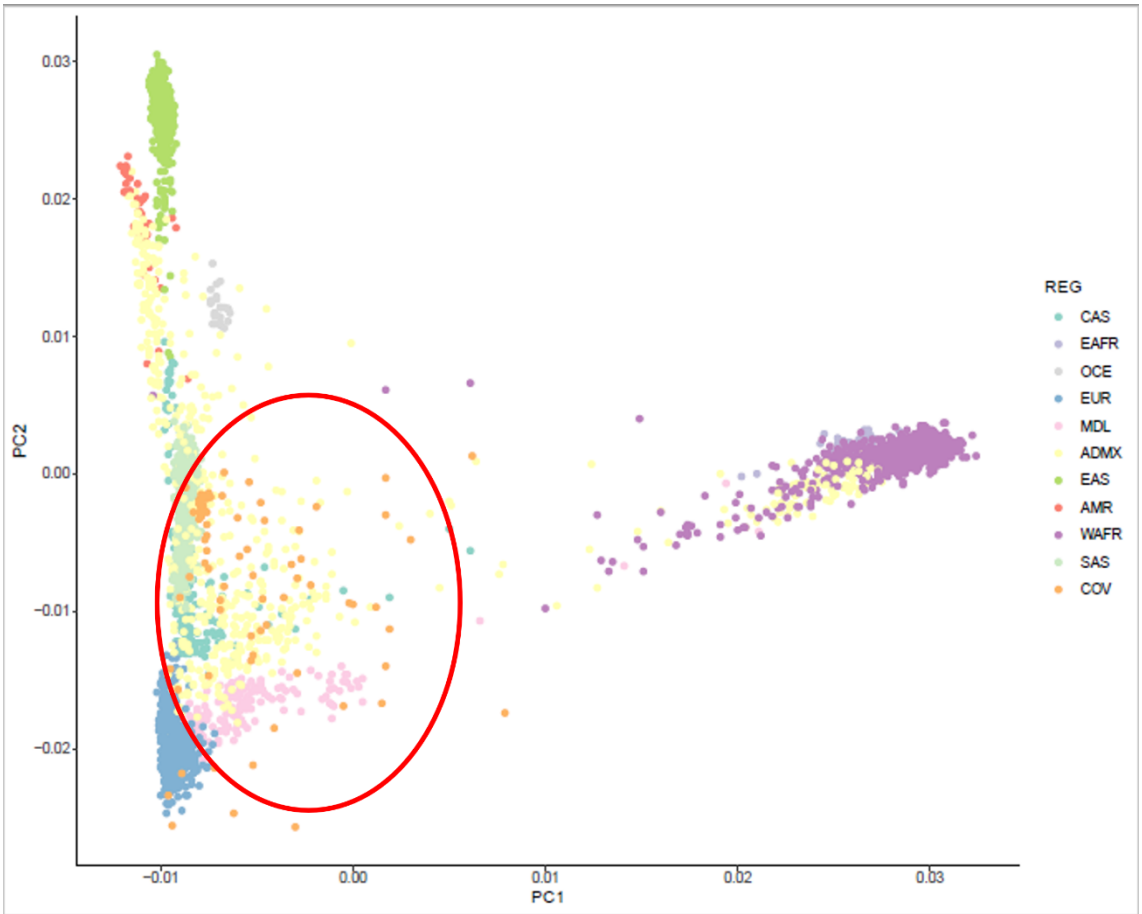


Figura 9. Análise de PCA das ancestralidades genômicas com base na frequência alélica presentes nos indivíduos do estudo e nos seguintes indivíduos: CAS - Ásia Central, EAFR - África Oriental, OCE - Oceania, EUR - Europa, MDL - Oriente Médio, ADMX - Populações Miscigenadas das Américas, EAS - Leste Asiático, AMR - Ameríndios, WAFR - África Ocidental, SAS - Sul/Sudeste Asiático, COV - Pacientes acometidos pela COVID-19.

Para avaliar se há associação entre a gravidade clínica da COVID-19 e ancestralidade genética, sexo e idade, foi realizado o modelo linear generalizado (GLM) binomial. Entretanto, não foi observado associação significativa entre gravidade e ancestralidade, apenas uma maior contribuição da ancestralidade dos povos originais nos

pacientes infectados pelo *SARS-CoV-2*, corroborando com os dados apresentados nos gráficos de análise supervisionada e de PCA. A única variável dependente no modelo a apresentar associação com a gravidade em COVID-19 foi a idade (p-valor = 0,0001).

Tabela 4. Média de ancestralidade em pacientes acometidos pela COVID-19. AFR - Africanos, EUR - Europeus, AMR - Ameríndios, alCOV - pacientes com a forma não grave da COVID-19, gCOV - pacientes com a forma grave da COVID-19.

	EUR	**AMR**	**AFR**
alCOV	0.2895046	0.6352316	0.07526385
gCOV	0.3273129	0.5916992	0.08098797
Todas	0.3039315	0.6186205	0.07744805

5.4 ANÁLISE DE ASSOCIAÇÃO GENÔMICA

A partir dos testes de associação para cada variante e levando em consideração a estruturação populacional e a idade como covariáveis, foram observadas nove variantes possivelmente associadas com à gravidade clínica da COVID-19 (Figura 10), das quais quatro não foram identificadas na literatura, e podem representar variantes exclusivas de populações da região.

Das variantes identificadas que estão associadas à gravidade clínica da COVID-19 temos: duas variantes sinônimas de baixo impacto rs40305 (*MYLK*) e rs469783 (*ERAP1*), uma *missense* de impacto moderado rs4840953 (*PRAG1*) e duas intrônicas de impacto modificador rs2839166 (*MCM3AP*) e rs12349255 (*NFIB*), observadas na Figura 10 e descritas na Tabela 5.

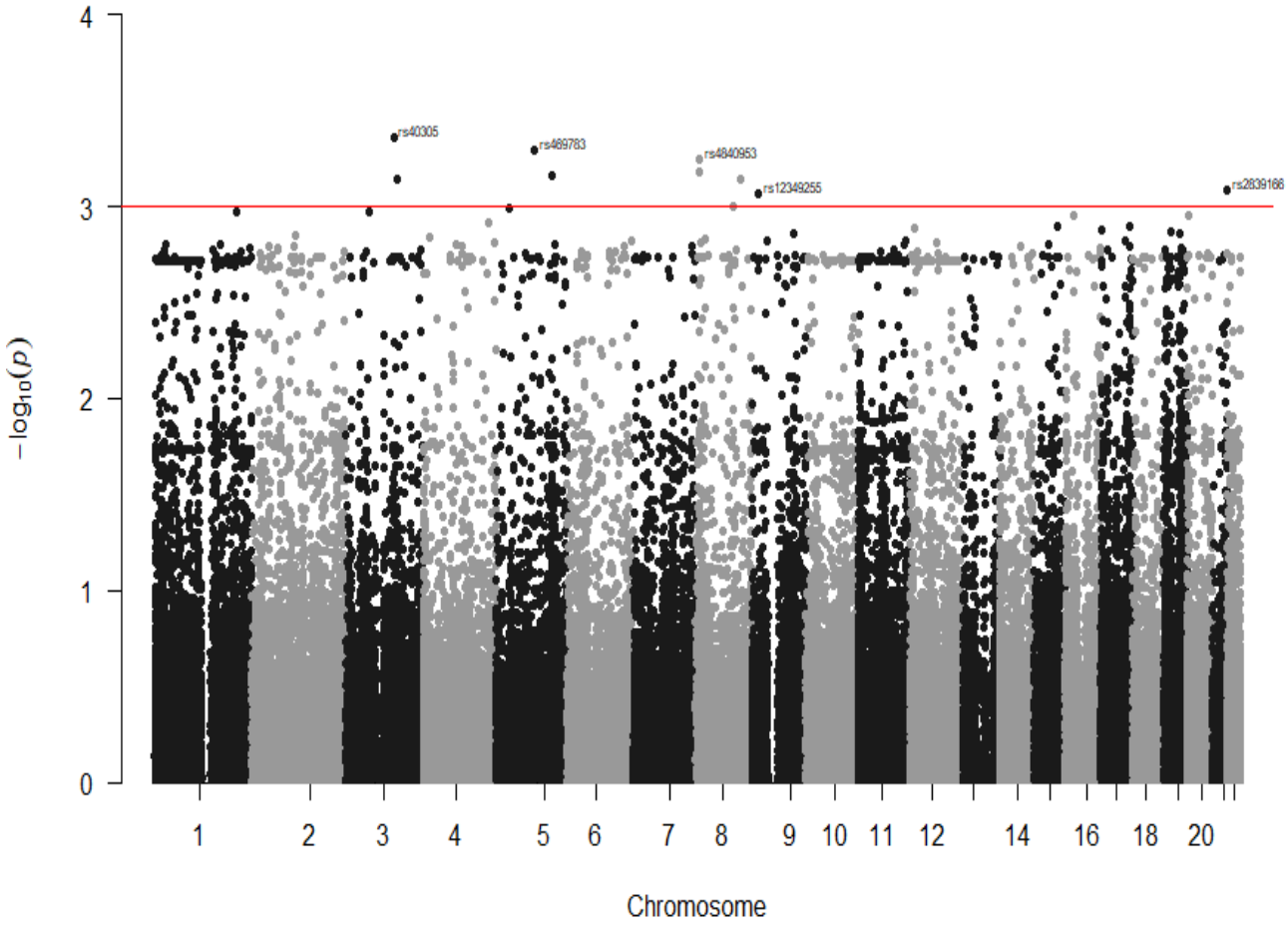


Figura 10. *Manhattan plot* ilustrando os polimorfismos presentes em indivíduos com a forma grave e não grave da COVID-19. A linha vermelha representa o *pvalor* de corte = 0.001.

Tabela 5. Polimorfismos associados significativamente (*pvalor* = 0.001) à gravidade clínica da COVID-19.

CHROM	POS	ID	REF	ALT	AA	TIPO	GENE	EFEITO	CLASSE	IMPACTO	PROCESSOS METABÓLICOS ASSOCIADOS AO GENE
chr3	123692742	rs40305	G	A	p.Thr1186Thr/c.3558C>T	SNV	MYLK	SYNONYMOUS_CODING	SILENCIOSA	BAIXO	Regulação positiva da migração celular; Morfogênese do tecido muscular da aorta
chr5	96785820	rs469783	C	T	p.Ala637Ala/c.1911G>A	SNV	ERAP1	SYNONYMOUS_CODING	SILENCIOSA	BAIXO	Resposta imune inata; Apresentação de antígeno na via MHC classe I
chr8	8376676	rs4840953	G	C	p.Ser578Cys/c.1733C>G	SNV	PRAG1	NON_SYNONYMOUS_CODING	MISSENSE	MODERADA	Migração celular
chr9	14155770	rs12349255	T	C	c.751+55A>G	SNV	NFIB	INTRON		MODIFICADOR	Replicação do DNA; Diferenciação celular
chr21	46236798	rs2839166	T	C	n.238-4036T>C	SNV	MCM3AP	INTRON		MODIFICADOR	Transporte de proteínas; Modificação de histonas

Das variantes observadas associadas, apenas uma variante (rs469783 C>T) já havia sido descrito na literatura associada a um fenótipo clínico. Desta forma, decidimos verificar e avaliar os processos metabólicos em que os genes em questão estão envolvidos.

A partir disso, é possível observar que as variantes possivelmente envolvidas na patogênese grave da COVID-19 apresentam um conjunto de mutações em genes associados a morfogênese do tecido muscular liso da aorta (*MYLK*), a resposta imune inata (*ERAP1*), a migração celular (*PRAG1*), a regulação da replicação do DNA (*NFIB*) e ao transporte de proteínas (*MCM3AP*) que podem estar associados a mecanismos de patogênese do *SARS-CoV-2* ainda não descritos.

6. DISCUSSÃO

O advento das tecnologias de sequenciamento de nova geração (NGS) permitiram o acesso a vasta quantidade de informações contidas nos genomas de diversos organismos, incluindo os de humanos. Nos últimos anos, grandes projetos como *HapMap* (†THE INTERNATIONAL HAPMAP CONSORTIUM, 2003), *1,000 Genomes Project* (THE 1000 GENOMES PROJECT CONSORTIUM et al., 2015), *Exac* (EXOME AGGREGATION CONSORTIUM et al., 2016) *e 100,000 Genomes Project* (NUNN et al., 2019), entre outras iniciativas pontuais têm se dedicado a: (i) caracterizar a variabilidade genética presente no genoma humano de diferentes grupos étnicos e geográficos; (ii) inovar os diagnósticos com base na genômica e serviços clínicos; (iii) avaliar novas ferramentas genômicas e de bioinformática para a prevenção e detecção precoce de doenças, com base nas pesquisas genômicas de ponta que tem gerado dados biológicos em grande escala.

As informações genômicas já geradas e armazenadas nesses bancos de dados mundiais, também têm contribuído para estudar a eficácia do uso do Sequenciamento Completo do Genoma (WGS) quer seja para acelerar o diagnóstico e o tratamento de doenças genéticas raras em recém-nascidos, que potencialmente podem levar as intervenções que salvam vidas; de doenças genéticas de manifestação tardia, que podem prevenir seu aparecimento e de doenças complexas; assim como no entendimento dos mecanismos farmacogenômicos (medicina de precisão) relacionados principalmente nos tratamentos de doenças crônicas, como doenças raras, câncer, e doenças infectocontagiosas, como COVID-19, Tuberculose, Hanseníase, entre outras (ZHAO & JONES & JONES, 2019; CASPAR et al., 2021).

No que diz respeito a população humana da Amazônia, esta é a menos representada do ponto de vista genômico, apesar de apresentar a maior diversidade biológica (alelos específicos e raros) distribuída entre povos tradicionais e miscigenados, que habitam áreas isoladas, rurais e urbanas (SCHAAN et al., 2020; RIBEIRO-DOS-SANTOS ET AL., 2020). Portanto, investigar a variabilidade genética individual e/ou populacional, pelo sequenciamento do exoma completo, pode auxiliar na resolução de muitas das questões, principalmente no que diz respeito a saúde desses povos (como os povos originais) que apresentam distintas respostas e vias metabólicas a fármacos, assim como risco elevado para muitas doenças genéticas, inflamatórias, infectocontagiosas (como a COVID-19) e resistência ou efeito adverso ao tratamento (RIBEIRO-DOS-SANTOS ET AL., 2013).

Somado a todas essas características descritas acima, desde a Idade Moderna, era raro observar doenças tão impactantes capazes de isolar populações humanas inteiras como um todo e em diversos âmbitos (socioeconômico, científico, sistemas de saúde), expondo de forma inequívoca as desigualdades existentes entre os diferentes países distribuídos pelo mundo, como foi a pandemia da COVID-19.

Esse conjunto de características envolvendo uma doença global, trouxe a necessidade de mudanças de estratégias, inovação, descoberta de novos mecanismos e de processos envolvidos no adoecimento populacional. Entretanto, apesar do crescente conhecimento sobre a biologia viral e fatores de risco da COVID-19, ainda perduram aqueles diversos aspectos sobre sua patogênese e progressão clínica, assim como dos fatores genéticos do hospedeiro que possam estar envolvidos.

Em concordância com resultados de fatores não genéticos associados a um pior prognóstico da COVID-19, obtidos em outros estudos (O'DRISCOLL et al., 2021; QUEIROZ et al., 2022; ÂNGULO-AGUADO et al., 2022), observou-se associação

significativa entre a idade (*pvalor* = 0.005; IC 95% = 8.98 – 21.2) e a presença de comorbidades (*pvalor* = 0.01; OR= 9.3; IC 95% = 2.97 – 33.05) no presente trabalho.

Dentre as diversas comorbidades que são associadas com a COVID-19, em nosso estudo, apenas a hipertensão arterial sistêmica (*pvalor* <0.01; OR= 10.1; IC 95% = 2.66 – 48.9) apresentou relevância significativa, provavelmente devido a hiperativação do sistema renina-angiotensina-aldosterona (RAAS) que aumenta a resposta inflamatória, o recrutamento de citocinas causando danos endoteliais (PAZ OCARANZA et al., 2020; RUIZ-STERNBERG et al., 2022).

Apesar de nenhuma das linhagens virais mais incidentes (B.1.1 *pvalor* = 0.42; OR= 1.6; IC 95% = 0.52 – 5.23; P.1 *pvalor* = 0.47; OR= 0.6; IC 95% = 0.23 – 1.89), terem sido associadas a gravidade da doença, já foi observado que a variante Gamma (P.1) pode estar relacionada a um maior número de mortes e de hospitalização de pacientes jovens no estado do Amazonas durante a segunda onda da COVID-19 em janeiro de 2021 (FREITAS et al., 2021).

Não foi observado associação significativa entre os sintomas mais prevalentes no grupo não grave [dor de garganta (43,7%), tosse (41,6%) e febre (22,9%)], corroborando com os resultados encontrados em outros estudos (DEL SER et al., 2021; IROUNGOU et al., 2021; RUIZ-STERNBERG et al., 2022).

No que diz respeito as características biológicas presentes nos pacientes acometidos com COVID-19, em relação às mutações patogênicas encontradas, houve uma prevalência considerável de mutações *missenses* (que alteram a leitura do aminoácido) não sinônimas nas regiões de éxon e sítios de *splicing*, com impacto moderado tanto no grupo grave (98,4%), quanto no grupo não grave (98,1%) (Figura 6). Além disso, foi observado uma maior proporção de mutações de alto impacto no grupo não grave quando comparado ao grupo grave (2,6%), que pode ser devido a proporção de

indivíduos distribuídos nos dois grupos (2,6%), uma vez que indivíduos que apresentaram a forma grave da COVID-19 formam menos de um terço do número amostral.

Adicionalmente, foram observadas 18 variantes de alto impacto e sete de substituição *framshift* em ambos os grupos, mas até o momento não existem estudos suficientes capazes de associá-las a um fenótipo clínico (Tabela 2). Portanto, nestas variantes avaliou-se as vias metabólicas envolvidas a esses genes.

Em relação às vias observadas no grupo grave, o gene *HADHA* está relacionado com a produção de proteínas mitocondriais e se faz necessário para a beta-oxidação de ácidos graxos nos cardiomiócitos (MIKLAS et al., 2019). As mutações observadas nesse gene podem ocasionar danos nos tecidos do coração, fígado e músculos, promovendo uma pior progressão clínica da COVID-19 (SYPNIEWSKI et al., 2022). Além disso, um relato de caso observou que um paciente com variantes patogênicas do *HADHA* desenvolveu insuficiência respiratória aguda, acompanhada de cardiomiopatia que o conduziu a uma falência múltipla de órgãos e ao seu pior prognóstico, o óbito desse indivíduo (WONGKITTICHOTE et al., 2020), corroborando com os achados deste estudo.

Outro gene observado no grupo grave, *MPZL2*, é membro da superfamília de imunoglobulinas (IgSF) expresso principalmente no pulmão e localizado no 11q23.3, locus gênico já associado à susceptibilidade ao câncer de pulmão (DAI et al., 2020; LIU et al., 2021). A partir disso, é possível sugerir que como o *MPZL2* atua na adesão das imunoglobulinas presentes no tecido pulmonar, mutações patogênicas que o impeçam de desenvolver suas funções básicas podem ocasionar danos na resposta imunológica a patógenos virais, por exemplo.

O gene *PALB2*, também observado no grupo grave, desempenha um papel crítico no reparo de recombinação homóloga (HRR) por meio de sua capacidade de recrutar

BRCA2 e *RAD51*, quando ocorrem quebras no material genético da molécula do DNA (YAMAMOTO & HIRASAWA, 2021). As mutações nesse gene já foram associadas ao desenvolvimento de diversos tipos de câncer como mama, pancreático e colorretal (XU et al., 2021). No presente estudo, a mutação observada no *PALB2* (NM_024675.4:c.503C>G, p.Ser168Ter) é preditada para interromper o funcionamento da proteína, causando perda de sua função na recombinação homóloga e prejudicando a via de reparo.

No que se refere aos genes observados no grupo não grave, tanto o gene *SAMD9* quanto o *ADA2* estão envolvidos na resposta imune inata do hospedeiro. O primeiro atua na defesa contra tipos de vírus, principalmente poxvírus, uma vez que codifica proteínas que podem afetar a replicação viral inibindo a proliferação celular. Entretanto esses mecanismos ainda permanecem incertos, sendo necessários mais estudos sobre a ação desse gene na resposta antiviral (XIE et al., 2021; PENG et al., 2022).

O gene *ADA2* promove a proliferação dos macrófagos, sendo já observado níveis elevados desse gene no plasma em indivíduos com leucemias, tuberculose e infecções pelo vírus HIV (LEE & AKSENTIJEVICH & ZHOU, 2022). Além disso, apesar de mutações nesse gene já serem associadas a fenótipos vasculares e inflamatórios, como AVC e vasculopatias sistêmicas (ZHOU et al., 2014), e ser um indicativo para desenvolver a forma grave da COVID-19, no presente estudo apenas um indivíduo carregava essa mutação em estado heterozigoto.

Além desses, também foi observado no grupo não grave o gene *NSMCE2*, componente do complexo SMC5/6, que participa dos processos de replicação, recombinação e reparo do DNA, além de silenciar vírus de DNA invasores como o HBV (vírus da hepatite B) (IRWAN & BOGERD & CULLEN, 2022). Ademais, Pond e colaboradores (2019) observaram que o *NSMCE2* previne danos mitóticos ao DNA

resultantes do estresse de replicação causado pelo tratamento com hidroxiureia (POND et al., 2019), a partir disso é possível hipotetizar que variantes nesse gene, mesmo patogênicas, podem ter efeito protetivo às infecções virais.

De acordo com as análises bioinformáticas, só há anotação da localização genômica de 72 das 1078 variantes missenses observadas, sendo principalmente intrônicas (55,5%) e exônicas (16,6%) (Tabela 3). Para o presente estudo, escolheu-se aprofundar as mutações missenses em regiões exônicas, visto que é o foco do trabalho.

Sendo assim, a variante rs1799945 (G>T) encontrada tanto no grupo grave quanto no não grave é presente no gene *HFE*, que controla o equilíbrio de ferro no organismo, e já foi investigada para a gravidade clínica da COVID-19 por Hubacek e colaboradores (2022), visto que o ferro é um catalisador essencial para processos metabólicos e uma sobrecarga de ferro é associada a uma pior progressão de infecções virais (SCHMIDT, 2020). Entretanto, essa variante ainda não foi associada significativamente com a gravidade da infecção pelo *SARS-CoV-2* (HUBACEK, 2021).

Além disso, a mutação C282Y presente no gene *HFE* já foi associada a condição hemocromatose hereditária (HH) que resulta no acúmulo de ferro nos órgãos podendo ocasionar danos progressivos em órgãos alvos. Indivíduos com HH apresentaram uma expressão reduzida de moléculas de MHC de classe I (RILEY et al., 2020), o que pode implicar em uma resposta imune precária do hospedeiro a infecção por patógenos.

O polimorfismo rs138291525 (G>T) ainda não foi associado a COVID-19, entretanto o gene *HOXA9*, no qual ele faz parte, já foi descrito como altamente expresso em leucemias pelo seu papel antiapoptótico ao manter as células cancerosas (MIYAMOTO et al., 2021). Como foi observado no grupo grave, essa variante pode estar associada a manutenção das células mesmo após a infecção pelo SARS-CoV-2 ou até mesmo ao escape do sistema imune do hospedeiro.

Kousathanas e colaboradores (2022) ao investigar o genoma de 7,491 indivíduos com a forma grave da COVID-19, observaram uma associação significativa entre a progressão da doença e o gene *FUT2*, responsável pela secreção de antígenos ABO que também já foram associados a um pior prognóstico da COVID-19 (COVID-19 HOST GENETICS INITIATIVE et al., 2021; KAUR & GUPTA & SAGAR, 2022). Além do mais, a variante encontrada por Kousathanas e colaboradores (2022) (rs368565 19:48697960) e as variantes encontradas no presente estudo (rs1800022 19:48703368 e rs200157007 19:4870329) são próximas genomicamente (KOUSATHANAS et al., 2022).

No que se refere as variantes encontradas no grupo não grave, o gene *ACAD11* (rs138611343), envolvido na beta-oxidação mitocondrial e no metabolismo de ácidos graxos, foi observado por Hammoudeh e colaboradores (2021) como *downregulated* no tecido hepático de pacientes com COVID-19 grave, provavelmente devido a ação do *SARS-CoV-2* no metabolismo do hospedeiro (SHEN et al., 2020; HAMMOUDEH et al., 2021).

As mutações presentes no gene PTCH1, frequentemente alterado no câncer de colorretal, contribuem na irregulação da via de sinalização Hedgehog que pode ocasionar na resistência a quimioterapia (HASANOVIC; MUS-VETEAU, 2018), todavia já foi observado também que este gene é associado a uma melhor sobrevivência por sensibilizar os tumores na imunoterapia (WANG et al., 2022), sendo assim, mais estudos são necessários a respeito da atuação desse gene.

Pela população brasileira ser altamente miscigenada e sofrer influência genética de diversos grupos populacionais, é essencial avaliar os padrões de estruturação populacional para controle de qualidade dos dados genéticos observados (RODRIGUES et al., 2019; RODRIGUES-SOARES et al., 2018) . Sendo assim, foi realizada uma análise

supervisionada no software Admixture utilizando as três ancestralidades mais importantes na formação da população brasileira: europeus, africanos e povos originais.

Na Figura 9 pode-se observar uma maior contribuição da ancestralidade de povos originais das Américas nos pacientes acometidos pela COVID-19, possivelmente devido à localização geográfica desses indivíduos (Região Norte, Amazônia do Brasil), onde residem grupos representativos desses povos. Adicionalmente, foi realizado o GLM binomial para investigar se há influência da ancestralidade genética na gravidade clínica da infecção do *SARS-CoV-2*, mas não foi encontrada associação significativa (Tabela 4).

Apesar disso, estudos como o de Shelton e colaboradores (2021) e Mathur e colaboradores (2021) já observaram diferenças na progressão clínica da COVID-19 de indivíduos de diferentes ancestralidades, como indivíduos com maior contribuição da ancestralidade afro-americana eram 82% mais propensos a serem hospitalizados e ancestralidades asiáticas e africanas possuem maior risco de hospitalização, ambos comparados com populações com maior ancestralidade europeia (MATHUR et al., 2021; SHELTON et al., 2021).

Entretanto, ainda são escassos os estudos que associam a ancestralidade genética individual à susceptibilidade ou gravidade da COVID-19 e buscam a distribuição genética de genes pré-selecionados em diferentes populações. Por conseguinte, Shen e colaboradores (2023) observaram que os genes *PPP2R2B* e o *LRRC4C* conferem uma maior susceptibilidade a infecção pelo *SARS-CoV-2* em indivíduos com maior contribuição européia e não são associados em indivíduos com maior contribuição africana (SHEN et al., 2023).

Além disso, Fabrizio e colaboradores (2022) analisaram polimorfismos presentes nos genes de risco *LZTFL1*, *SLC6A20*, *MHC*, *ABO*, *DPP9* e *IFNAR2* e observaram que populações predominantemente africanas possuem menos risco de infecção ao *SARS-*

CoV-2 do que populações asiáticas e europeias (FABRIZIO et al., 2022). O fato de os estudos serem discordantes reforça a importância de analisar a ancestralidade genômica de cada população para uma melhor compreensão do prognóstico da doença e adaptação das estratégias de prevenção e tratamento de cada grupo populacional.

Foram observados cinco variantes possivelmente associadas à gravidade clínica da COVID-19: rs40305 (*MYLK*), rs469783 (*ERAP1*), rs4840953 (*PRAG1*), rs2839166 (*MCM3AP*) e rs12349255 (*NFIB*) observadas na Figura 10 e descritas na Tabela 5. Além dessas, foram observadas quatro variantes sem identificação prévia na literatura e por conta disso podem ser variantes exclusivas de populações da região. Essas variantes não identificadas podem refletir a falta de estudos e de informações genômicas nas populações que habitam a região.

Por outro lado, como na literatura só há dados de apenas uma das variantes (rs469783 C>T) aqui observada, decidiu-se investigar como os genes poderiam estar envolvidos e modificando as vias metabólicas para o desenvolvimento da forma grave da COVID-19.

A variante rs469783 (C>T) foi a única encontrada neste estudo já associada a um fenótipo clínico na literatura, sendo observada por Yang e colaboradores (2021) como associada a sobrevivência de indivíduos de origem europeia com câncer de pulmão de células não-pequenas (YANG et al., 2021).

Investigações genéticas anteriores já demonstraram que haplótipos do gene *MYLK*, um regulador da permeabilidade celular na resposta a inflamação, pode conferir suscetibilidade à asma severa e à sepse induzida pela Síndrome Respiratória Aguda Grave (SZILÁGYI et al., 2017). Os mecanismos patogênicos associados ao aumento da permeabilidade vascular durante a infecção pelo *SARS-CoV-2* permanecem

incompreendidos, entretanto podem estar relacionados ao citoesqueleto das células endoteliais, essenciais para a regulação e reparo da barreira vascular (SUN et al., 2021).

A variante presente no gene *ERAP1*, realiza a trimagem final de peptídeos de ligação da classe I do HLA (Antígeno Leucocitário Humano) onde ocorre o reconhecimento e eliminação de células infectadas por vírus pelo sistema imune (PEPELYAYEVA; AMALFITANO, 2019). Considerando este achado, o *ERAP1* é envolvido na regulação da sinalização pró-inflamatória de citocinas (FAN et al., 2022) e na regulação do RAS (Sistema Renina-Angiotensina-Aldosterona) (D'AMICO et al., 2021), ambos mecanismos cruciais no desenvolvimento da forma grave da COVID-19. Além disso, já foi demonstrado que o gene *ERAP1* aumenta a atividade fagocítica dos macrófagos em resposta ao interferon-gamma por um mecanismo dependente de receptores TLR, justamente um dos mecanismos de resposta do hospedeiro a infecção pelo *SARS-CoV-2* (LÓPEZ DE CASTRO, 2018; MERAD et al., 2022).

Outra variante também associada à gravidade da COVID-19 é a rs4840953 (G>C) do gene *PRAG1*, esta única variante *missense* encontrada, pertence a subfamília de pseudo-kinases e está envolvida na adesão e crescimento celular (LECOINTRE et al., 2018). Devido a isso, já foi encontrada superexpressa em células de adenocarcinoma por promover a migração e invasão celular (TACTACAN et al., 2015).

O gene *NFIB* também já foi associado ao câncer por possível promoção de mudanças no estado da cromatina em células cancerosas e assim facilitar a migração, invasão e metástase (CAMPBELL et al., 2018; XU et al., 2019), todavia mais estudos precisam ser realizados para confirmação da sua atuação.

Por último, o gene *MCM3AP* codifica a proteína GANP, fundamental na exportação de mRNA do núcleo para o citoplasma (DONG et al., 2020). A partir disso é possível sugerir que variantes presentes no gene *MCM3AP*, como a rs2839166 encontrada

no estudo, possam estar envolvidas no transporte do *SARS-CoV-2* do núcleo para o citoplasma celular, mas tal mecanismo deve ser estudado mais profundamente.

Portanto, é essencial a investigação da relação de variantes em genes com os mecanismos de patogênese da forma grave da COVID-19, principalmente utilizando plataformas de sequenciamento massivo do exoma e das análises de bioinformáticas, especialmente em populações que habitam a Amazônia, uma das regiões menos investigadas do ponto de vista genômico, mas a de maior representatividade de povos tradicionais (povos originários e afrodescendentes) e consequentemente a como maior número de alelos raros ainda não descritos na literatura mundial.

7. CONCLUSÃO

A COVID-19 difere das demais doenças pelo seu amplo espectro clínico e sistêmico, englobando desde casos assintomáticos ou leves até casos graves de insuficiência respiratória e óbito. Esta variação de sintomas clínicos pode ser decorrente das diferentes linhagens virais, mas também pode ocorrer em indivíduos infectados pela mesma linhagem, sugerindo que há outros fatores envolvidos no agravamento da doença, além de fatores socioeconômicos, das diferentes políticas de combate à COVID-19 adotadas por cada país e de fatores não genéticos do hospedeiro, como idade, sexo e comorbidades prévias, já associados ao pior prognóstico da COVID-19.

Com o objetivo de caracterizar marcadores genômicos associados à infecção por *SARS-CoV*-2, assim como sua gravidade clínica, foi realizado o sequenciamento total do exoma de indivíduos acometidos pela COVID-19 e daqueles que apresentaram a forma não grave e a forma grave dessa doença. Assim como em outros estudos, também foi encontrado associação entre a idade (pvalor = 0.005; IC 95% = 8.98 – 21.2), linhagem viral (pvalor = 0.05) e presença de comorbidades (pvalor = 0.01; OR= 9.3; IC 95% = 2.97 – 33.05) a gravidade clínica da COVID-19. Não foram observadas variantes genômicas de risco à infecção pelo *SARS-CoV-2*. Por outro lado, foram observadas nove (9) variantes possivelmente associadas à gravidade clínica da COVID-19, entre as quais quatro não se encontram nenhuma anotação genômica e possivelmente refletem mutações específicas das população da Amazônia, e as outras cinco são rs40305 (*MYLK*), rs469783 (*ERAP1*), rs4840953 (*PRAG1*), rs2839166 (*MCM3A*P) e rs12349255 (*NFIB*). De forma independente, nenhuma dessas mutações já havia sido descrita previamente associada a COVID-19 e por isso necessitam-se de investigações futuras para confirmação.

Até onde sabemos, esse é o primeiro estudo de sequenciamento total do exoma em indivíduos residentes no estado do Pará que busca investigar variantes genômicas presentes em pacientes acometidos pela COVID-19, associados aos quadros de gravidade clínica da COVID-19, reforçando a importância de se realizar mais estudos genômicos em populações que habitam a região Amazônica, uma das mais diversas do ponto de vista da presença de alelos raros e específicos. Esse estudo faz parte de um projeto maior onde se pretende realizar o exoma de 300 indivíduos que foram acometidos pela COVID-19, a fim de se associar marcadores genômicos com o risco de desenvolvimento da doença e com a gravidade clínica pela infecção pelo *SARS-CoV-2*.

REFERÊNCIAS

AMOROSO, A. et al. HLA and AB0 Polymorphisms May Influence SARS-CoV-2 Infection and COVID-19 Severity. **Transplantation**, v. 105, n. 1, p. 193–200, jan. 2021.

ANASTASSOPOULOU, C. et al. Human genetic factors associated with susceptibility to SARS-CoV-2 infection and COVID-19 disease severity. **Human Genomics**, v. 14, n. 1, p. 40, dez. 2020.

ANDOLFO, I. et al. Common variants at 21q22.3 locus influence MX1 and TMPRSS2 gene expression and susceptibility to severe COVID-19. **iScience**, v. 24, n. 4, p. 102322, abr. 2021.

ANGULO-AGUADO, M. et al. Association Between the LZTFL1 rs11385942 Polymorphism and COVID-19 Severity in Colombian Population. **Frontiers in Medicine**, v. 9, p. 910098, 20 jun. 2022.

ARAUJO, D. B. et al. SARS-CoV-2 isolation from the first reported patients in Brazil and establishment of a coordinated task network. **Memórias do Instituto Oswaldo Cruz**, v. 115, p. e200342, 2020.

BALDASSARRI, M. et al. Severe COVID-19 in Hospitalized Carriers of Single CFTR Pathogenic Variants. **Journal of Personalized Medicine**, v. 11, n. 6, p. 558, 15 jun. 2021.

BAMSHAD, M. J. et al. Exome sequencing as a tool for Mendelian disease gene discovery. **Nature Reviews Genetics**, v. 12, n. 11, p. 745–755, nov. 2011.

BEALS, J. K. GWAS vs Whole-Exome Sequencing: What's the Difference and Why We Should Care. p. 2, [s.d.].

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 1 ago. 2014.

BOSON, B. et al. The SARS-CoV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles. **Journal of Biological Chemistry**, v. 296, p. 100111, jan. 2021.

BRODIN, P. Immune determinants of COVID-19 disease presentation and severity. **Nature Medicine**, v. 27, n. 1, p. 28–33, jan. 2021.

CAMPBELL, T. M. et al. ERα Binding by Transcription Factors NFIB and YBX1 Enables FGFR2 Signaling to Modulate Estrogen Responsiveness in Breast Cancer. **Cancer Research**, v. 78, n. 2, p. 410–421, 15 jan. 2018.

CANTALUPO, S. et al. Regulatory Noncoding and Predicted Pathogenic Coding Variants of CCR5 Predispose to Severe COVID-19. **International Journal of Molecular Sciences**, v. 22, n. 10, p. 5372, 20 maio 2021.

CASPAR SM, SCHNEIDER T, STOLL P, MEIENBERG J, MATYAS G. Potential of whole-genome sequencing-based pharmacogenetic profiling. Pharmacogenomics. 2021 Feb;22(3):177-190. doi: 10.2217/pgs-2020-0155. Epub 2021 Feb 1. PMID: 33517770.

CASTELLI, E. C. et al. MHC Variants Associated With Symptomatic Versus Asymptomatic SARS-CoV-2 Infection in Highly Exposed Individuals. **Frontiers in Immunology**, v. 12, p. 742881, 28 set. 2021.

CHAMNANPHON, M. et al. Host genetic factors of COVID-19 susceptibility and disease severity in a Thai population. **Journal of Human Genetics**, v. 67, n. 5, p. 295–301, maio 2022.

CHAN, J. F.-W. et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. **Emerging Microbes & Infections**, v. 9, n. 1, p. 221–236, 1 jan. 2020.

CHANDLER, N. et al. Rapid prenatal diagnosis using targeted exome sequencing: a cohort study to assess feasibility and potential impact on prenatal counseling and pregnancy management. **Genetics in Medicine**, v. 20, n. 11, p. 1430–1437, nov. 2018.

CHEN, G. et al. Clinical and immunological features of severe and moderate coronavirus disease 2019. **Journal of Clinical Investigation**, v. 130, n. 5, p. 2620–2629, 13 abr. 2020.

CHOUDHARY, S. et al. Role of Genetic Variants and Gene Expression in the Susceptibility and Severity of COVID-19. **Annals of Laboratory Medicine**, v. 41, n. 2, p. 129–138, 1 mar. 2021.

CINGOLANI, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w ^{1118} ; iso-2; iso-3. **Fly**, v. 6, n. 2, p. 80–92, abr. 2012.

COVID-19 E OS POVOS INDÍGENAS. **COVID-19 e os Povos Indígenas**. Disponível em: <https://covid19.socioambiental.org/>. Acesso em: 20 jul. 2022.

COVID-19 HOST GENETICS INITIATIVE et al. Mapping the human genetic architecture of COVID-19. **Nature**, v. 600, n. 7889, p. 472–477, 16 dez. 2021.

D'AMICO, S. et al. ERAP1 and ERAP2 Enzymes: A Protective Shield for RAS against COVID-19? **International Journal of Molecular Sciences**, v. 22, n. 4, p. 1705, 8 fev. 2021.

D'ANTONIO, M. et al. SARS-CoV-2 susceptibility and COVID-19 disease severity are associated with genetic variants affecting gene expression in a variety of tissues. **Cell Reports**, v. 37, n. 7, p. 110020, nov. 2021.

DAI, J. et al. Genome-wide association study of INDELs identified four novel susceptibility loci associated with lung cancer risk. **International Journal of Cancer**, v. 146, n. 10, p. 2855–2864, 15 maio 2020.

DEHGHAN, A. Genome-Wide Association Studies. **Genetic Epidemiology**, Methods in Molecular Biology. v. 1793, p. 37–49, 2018.

DEL SER, T. et al. Residence, Clinical Features, and Genetic Risk Factors Associated with Symptoms of COVID-19 in a Cohort of Older People in Madrid. **Gerontology**, v. 67, n. 3, p. 281–289, 2021.

DEVAUX, C. A.; ROLAIN, J.-M.; RAOULT, D. ACE2 receptor polymorphism: Susceptibility to SARS-CoV-2, hypertension, multi-organ failure, and COVID-19 disease outcome. **Journal of Microbiology, Immunology and Infection**, v. 53, n. 3, p. 425–435, jun. 2020.

DOLLED-FILHART, M. P. et al. Computational and Bioinformatics Frameworks for Next-Generation Whole Exome and Genome Sequencing. **The Scientific World Journal**, v. 2013, p. 1–10, 2013.

DONG, H. et al. Genetic spectrum of *MCM3AP* and its relationship with phenotype of CHARCOT-MARIE-TOOTH disease. **Journal of the Peripheral Nervous System**, v. 25, n. 2, p. 107–111, jun. 2020.

DOWNES, D. J. et al. Identification of LZTFL1 as a candidate effector gene at a COVID-19 risk locus. **Nature Genetics**, v. 53, n. 11, p. 1606–1615, nov. 2021.

ELLINGFORD, J. M. et al. Validation of copy number variation analysis for next-generation sequencing diagnostics. **European Journal of Human Genetics**, v. 25, n. 6, p. 719–724, jun. 2017.

ENOCH, M.-A. et al. Using ancestry-informative markers to define populations and detect population stratification. **Journal of Psychopharmacology**, v. 20, n. 4_suppl, p. 19–26, jul. 2006.

EXOME AGGREGATION CONSORTIUM et al. Analysis of protein-coding genetic variation in 60,706 humans. **Nature**, v. 536, n. 7616, p. 285–291, ago. 2016.

FABRIZIO, C. et al. Analysis of Genetic Variants Associated with COVID-19 Outcome Highlights Different Distributions among Populations. **Journal of Personalized Medicine**, v. 12, n. 11, p. 1851, 5 nov. 2022.

FAIRLEY, S. et al. The International Genome Sample Resource (IGSR) collection of open human genomic variation resources. **Nucleic Acids Research**, v. 48, n. D1, p. D941–D947, 8 jan. 2020.

FAN, J. et al. Genome-wide association study of SARS-CoV-2 infection in Chinese population. **European Journal of Clinical Microbiology & Infectious Diseases**, 4 ago. 2022.

FARWELL, K. D. et al. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model–based analysis: results from 500 unselected families with undiagnosed genetic conditions. **Genetics in Medicine**, v. 17, n. 7, p. 578–586, jul. 2015.

FINDLAY, G. M. et al. Accurate classification of BRCA1 variants with saturation genome editing. **Nature**, v. 562, n. 7726, p. 217–222, out. 2018.

FITZGERALD, K. A.; KAGAN, J. C. Toll-like Receptors and the Control of Immunity. **Cell**, v. 180, n. 6, p. 1044–1066, mar. 2020.

FREITAS, A. R. R. et al. The emergence of novel SARS-CoV-2 variant P.1 in Amazonas (Brazil) was temporally associated with a change in the age and sex profile of COVID-

19 mortality: A population based ecological study. **The Lancet Regional Health - Americas**, v. 1, p. 100021, set. 2021.

GILISSEN, C. et al. Disease gene identification strategies for exome sequencing. **European Journal of Human Genetics**, v. 20, n. 5, p. 490–497, maio 2012.

GUERRA, F. P. et al. Exome resequencing and GWAS for growth, ecophysiology, and chemical and metabolomic composition of wood of Populus trichocarpa. **BMC Genomics**, v. 20, n. 1, p. 875, dez. 2019.

HALLAL, P. C. et al. SARS-CoV-2 antibody prevalence in Brazil: results from two successive nationwide serological household surveys. **The Lancet Global Health**, v. 8, n. 11, p. e1390–e1398, nov. 2020.

HAMMOUDEH, S. M. et al. Insight into molecular mechanisms underlying hepatic dysfunction in severe COVID-19 patients using systems biology. **World Journal of Gastroenterology**, v. 27, n. 21, p. 2850–2870, 7 jun. 2021.

HASANOVIC, A.; MUS-VETEAU, I. Targeting the Multidrug Transporter Ptch1 Potentiates Chemotherapy Efficiency. **Cells**, v. 7, n. 8, p. 107, 14 ago. 2018.

HEAD, S. R. et al. Library construction for next-generation sequencing: Overviews and challenges. **BioTechniques**, v. 56, n. 2, p. 61–77, fev. 2014.

HOFFMANN, M. et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. **Cell**, v. 181, n. 2, p. 271- 280.e8, abr. 2020.

HUANG, S.-W.; WANG, S.-F. SARS-CoV-2 Entry Related Viral and Host Genetic Variations: Implications on COVID-19 Severity, Immune Escape, and Infectivity. **International Journal of Molecular Sciences**, v. 22, n. 6, p. 3060, 17 mar. 2021.

HUBACEK, J. Effects of Selected Inherited Factors on Susceptibility to SARS-CoV-2 Infection and COVID-19 Progression. **Physiological Research**, p. s125-S134, 14 dez. 2021.

IROUNGOU, B. A. et al. Demographic and Clinical Characteristics Associated With Severity, Clinical Outcomes, and Mortality of COVID-19 Infection in Gabon. **JAMA Network Open**, v. 4, n. 9, p. e2124190, 14 set. 2021.

IRWAN, I. D.; BOGERD, H. P.; CULLEN, B. R. Epigenetic silencing by the SMC5/6 complex mediates HIV-1 latency. **Nature Microbiology**, v. 7, n. 12, p. 2101–2113, 14 nov. 2022.

ISHAK, A. et al. The association of COVID-19 severity and susceptibility and genetic risk factors: A systematic review of the literature. **Gene**, v. 836, p. 146674, ago. 2022.

IZCOVICH, A. et al. Prognostic factors for severity and mortality in patients infected with COVID-19: A systematic review. **PLOS ONE**, v. 15, n. 11, p. e0241955, 17 nov. 2020.

JACKSON, C. B. et al. Mechanisms of SARS-CoV-2 entry into cells. **Nature Reviews Molecular Cell Biology**, v. 23, n. 1, p. 3–20, jan. 2022.

JIN, Y. et al. Virology, Epidemiology, Pathogenesis, and Control of COVID-19. **Viruses**, v. 12, n. 4, p. 372, 27 mar. 2020.

KARCZEWSKI, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. **Nature**, v. 581, n. 7809, p. 434–443, 28 maio 2020.

KAUR, P.; GUPTA, M.; SAGAR, V. FUT2 gene as a genetic susceptible marker of infectious diseases: A Review. **International Journal of Molecular Epidemiology and Genetics**, v. 13, n. 1, p. 1–14, 2022.

KHALIL, B. A.; ELEMAM, N. M.; MAGHAZACHI, A. A. Chemokines and chemokine receptors during COVID-19 infection. **Computational and Structural Biotechnology Journal**, v. 19, p. 976–988, 2021.

KHANNA, R. et al. COVID-19 pandemic: Lessons learned and future directions. **Indian Journal of Ophthalmology**, v. 68, n. 5, p. 703, 2020.

KHARE, S. et al. GISAID's Role in Pandemic Response. **China CDC Weekly**, v. 3, n. 49, p. 1049–1051, 2021.

KHAYAT, A. S. et al. ACE2 polymorphisms as potential players in COVID-19 outcome. **PLOS ONE**, v. 15, n. 12, p. e0243887, 28 dez. 2020.

KIMURA, L.; LEMES, R. B.; NUNES, K. Ancestralidade: genética, herança e identidade. **Genética na Escola**, v. 17, n. 1, p. 41–52, 16 mar. 2022.

KOBOLDT, D. C. Best practices for variant calling in clinical sequencing. **Genome Medicine**, v. 12, n. 1, p. 91, dez. 2020.

KOSMICKI, J. A. et al. Pan-ancestry exome-wide association analyses of COVID-19 outcomes in 586,157 individuals. **The American Journal of Human Genetics**, v. 108, n. 7, p. 1350–1355, jul. 2021.

KOUSATHANAS, A. et al. Whole-genome sequencing reveals host factors underlying critical COVID-19. **Nature**, v. 607, n. 7917, p. 97–103, 7 jul. 2022.

LADUCA, H. et al. Exome sequencing covers >98% of mutations identified on targeted next generation sequencing panels. **PLOS ONE**, v. 12, n. 2, p. e0170843, 2 fev. 2017.

LANDRUM, M. J. et al. ClinVar: improving access to variant interpretations and supporting evidence. **Nucleic Acids Research**, v. 46, n. D1, p. D1062–D1067, 4 jan. 2018.

LECOINTRE, C. et al. Dimerization of the Pragmin Pseudo-Kinase Regulates Protein Tyrosine Phosphorylation. **Structure**, v. 26, n. 4, p. 545- 554.e4, abr. 2018.

LEE, P. Y.; AKSENTIJEVICH, I.; ZHOU, Q. Mechanisms of vascular inflammation in deficiency of adenosine deaminase 2 (DADA2). **Seminars in Immunopathology**, v. 44, n. 3, p. 269–280, maio 2022.

LEE, S. J.; KIM, Y.-J.; AHN, D.-G. Distinct Molecular Mechanisms Characterizing Pathogenesis of SARS-CoV-2. **Journal of Microbiology and Biotechnology**, v. 32, n. 9, p. 1073–1085, 28 set. 2022.

LEI, T. et al. Whole-exome sequencing for prenatal diagnosis of fetuses with congenital anomalies of the kidney and urinary tract. **Nephrology Dialysis Transplantation**, v. 32, n. 10, p. 1665–1675, 1 out. 2017.

LEIPZIG, J. A review of bioinformatic pipeline frameworks. **Briefings in Bioinformatics**, p. bbw020, 24 mar. 2016.

LI, H. **Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM**. arXiv, , 26 maio 2013. Disponível em: <http://arxiv.org/abs/1303.3997>. Acesso em: 1 set. 2022

LI, H. et al. The Sequence Alignment/Map format and SAMtools. **Bioinformatics**, v. 25, n. 16, p. 2078–2079, 15 ago. 2009.

LI, H.; DURBIN, R. Fast and accurate short read alignment with Burrows-Wheeler transform. **Bioinformatics**, v. 25, n. 14, p. 1754–1760, 15 jul. 2009.

LIU, Y. et al. Rare deleterious germline variants and risk of lung cancer. **npj Precision Oncology**, v. 5, n. 1, p. 12, 16 fev. 2021.

LIU, Y.-C.; KUO, R.-L.; SHIH, S.-R. COVID-19: The first documented coronavirus pandemic in history. **Biomedical Journal**, v. 43, n. 4, p. 328–333, ago. 2020.

LÓPEZ DE CASTRO, J. A. How ERAP1 and ERAP2 Shape the Peptidomes of Disease-Associated MHC-I Proteins. **Frontiers in Immunology**, v. 9, p. 2463, 30 out. 2018.

MALAQUIAS, M. A. S. et al. The role of the lectin pathway of the complement system in SARS-CoV-2 lung injury. **Translational Research**, v. 231, p. 55–63, maio 2021.

MARIANO, G. et al. Structural Characterization of SARS-CoV-2: Where We Are, and Where We Need to Be. **Frontiers in Molecular Biosciences**, v. 7, p. 605236, 17 dez. 2020.

MARTIN, A. R. et al. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. **The American Journal of Human Genetics**, v. 100, n. 4, p. 635–649, abr. 2017.

MATHUR, R. et al. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. **The Lancet**, v. 397, n. 10286, p. 1711–1724, maio 2021.

MENG, L. et al. Use of Exome Sequencing for Infants in Intensive Care Units: Ascertainment of Severe Single-Gene Disorders and Effect on Medical Management. **JAMA Pediatrics**, v. 171, n. 12, p. e173438, 4 dez. 2017.

MERAD, M. et al. The immunology and immunopathology of COVID-19. **Science**, v. 375, n. 6585, p. 1122–1127, 11 mar. 2022.

MIKLAS, J. W. et al. TFPa/HADHA is required for fatty acid beta-oxidation and cardiolipin re-modeling in human cardiomyocytes. **Nature Communications**, v. 10, n. 1, p. 4671, 11 out. 2019.

MINISTÉRIO DA SAÚDE. **Painel Coronavírus**. Disponível em: <https://covid.saude.gov.br/>. Acesso em: 4 fev. 2023.

MIYAMOTO, R. et al. HOXA9 promotes MYC-mediated leukemogenesis by maintaining gene expression for multiple anti-apoptotic pathways. **eLife**, v. 10, p. e64148, 26 jul. 2021.

MOONESINGHE, R. et al. Estimating the contribution of genetic variants to difference in incidence of disease between population groups. **European Journal of Human Genetics**, v. 20, n. 8, p. 831–836, ago. 2012.

NAKANISHI, T. et al. Age-dependent impact of the major common genetic risk factor for COVID-19 on severity and mortality. **Journal of Clinical Investigation**, v. 131, n. 23, p. e152386, 1 dez. 2021.

NASLAVSKY, M. S. et al. Exomic variants of an elderly cohort of Brazilians in the ABraOM database. **Human Mutation**, v. 38, n. 7, p. 751–763, jul. 2017.

NG, S. B. et al. Exome sequencing identifies the cause of a mendelian disorder. **Nature Genetics**, v. 42, n. 1, p. 30–35, jan. 2010.

NGUYEN, A. et al. Human Leukocyte Antigen Susceptibility Map for Severe Acute Respiratory Syndrome Coronavirus 2. **Journal of Virology**, v. 94, n. 13, p. e00510-20, 16 jun. 2020.

O'DRISCOLL, M. et al. Age-specific mortality and immunity patterns of SARS-CoV-2. **Nature**, v. 590, n. 7844, p. 140–145, 4 fev. 2021.

OCHANI, R. K. et al. COVID-19 pandemic: from origins to outcomes. A comprehensive review of viral pathogenesis, clinical manifestations, diagnostic evaluation, and management. p. 17, 2021.

ONYEAKA, H. et al. COVID-19 pandemic: A review of the global lockdown and its far-reaching effects. **Science Progress**, v. 104, n. 2, p. 003685042110198, abr. 2021.

PAIRO-CASTINEIRA, E. et al. Genetic mechanisms of critical illness in COVID-19. **Nature**, v. 591, n. 7848, p. 92–98, 4 mar. 2021.

PANDIT, R. et al. First report on genome wide association study in western Indian population reveals host genetic factors for COVID-19 severity and outcome. **Genomics**, v. 114, n. 4, p. 110399, jul. 2022.

PAZ OCARANZA, M. et al. Counter-regulatory renin–angiotensin system in cardiovascular disease. **Nature Reviews Cardiology**, v. 17, n. 2, p. 116–129, fev. 2020.

PENG, S. et al. Structure and function of an effector domain in antiviral factors and tumor suppressors SAMD9 and SAMD9L. **Proceedings of the National Academy of Sciences**, v. 119, n. 4, p. e2116550119, 25 jan. 2022.

PEPELYAYEVA, Y.; AMALFITANO, A. The role of ERAP1 in autoinflammation and autoimmunity. **Human Immunology**, v. 80, n. 5, p. 302–309, maio 2019.

PEREIRA, A. C. et al. Genetic risk factors and COVID-19 severity in Brazil: results from BRACOVID study. **Human Molecular Genetics**, p. ddac045, 2 abr. 2022.

PEREIRA, H. DOS S. et al. Fases da pandemia em 6 capitais da Amazônia brasileira. **Fases da pandemia em 6 capitais da Amazônia brasileira**, v. 2, n. 11, jun. 2020.

PETROVSKI, S. et al. Whole-exome sequencing in the evaluation of fetal structural anomalies: a prospective cohort study. **The Lancet**, v. 393, n. 10173, p. 758–767, fev. 2019.

PLENGE, R. M. Molecular Underpinnings of Severe Coronavirus Disease 2019. **JAMA**, v. 324, n. 7, p. 638, 18 ago. 2020.

POND, K. W. et al. Rescue of collapsed replication forks is dependent on NSMCE2 to prevent mitotic DNA damage. **PLOS Genetics**, v. 15, n. 2, p. e1007942, 8 fev. 2019.

QUEIROZ, M. A. F. et al. Cytokine Profiles Associated With Acute COVID-19 and Long COVID-19 Syndrome. **Frontiers in Cellular and Infection Microbiology**, v. 12, p. 922422, 30 jun. 2022.

RAHMAN, S. et al. Epidemiology, pathogenesis, clinical presentations, diagnosis and treatment of COVID-19: a review of current evidence. **Expert Review of Clinical Pharmacology**, v. 14, n. 5, p. 601–621, 4 maio 2021.

RAMAMOORTHY, A. et al. Racial and Ethnic Differences in Drug Disposition and Response: Review of New Molecular Entities Approved Between 2014 and 2019. **The Journal of Clinical Pharmacology**, v. 62, n. 4, p. 486–493, abr. 2022.

RAVI, V.; SAXENA, S.; PANDA, P. S. Basic virology of SARS-CoV 2. **Indian Journal of Medical Microbiology**, v. 40, n. 2, p. 182–186, abr. 2022.

REHM, H. L. et al. ACMG clinical laboratory standards for next-generation sequencing. **Genetics in Medicine**, v. 15, n. 9, p. 733–747, set. 2013.

RIBEIRO-DOS-SANTOS AM, DE SOUZA JES, ALMEIDA R, ALENCAR DO, BARBOSA MS, et al. (2013) High-Throughput Sequencing of a South American Amerindian. PLOS ONE 8(12): e83340.

RIBEIRO-DOS-SANTOS, A. M., VIDAL, A. F., VINASCO-SANDOVAL, T., GUERREIRO, J., SANTOS, S., RIBEIRO-DOS-SANTOS, Â., & DE SOUZA, S. J. (2020). Exome sequencing of native populations from the Amazon reveals patterns on the peopling of South America. Frontiers in Genetics, 11, 548507.

RILEY, M. J. et al. Hereditary haemochromatosis, haemophagocytic lymphohistiocytosis and COVID-19. **Clinical Infection in Practice**, v. 7–8, p. 100052, out. 2020.

RODRIGUES, J. C. G. et al. Polymorphisms of ADME-related genes and their implications for drug safety and efficacy in Amazonian Amerindians. **Scientific Reports**, v. 9, n. 1, p. 7201, dez. 2019.

RODRIGUES-SOARES, F. et al. Genetic structure of pharmacogenetic biomarkers in Brazil inferred from a systematic review and population-based cohorts: a RIBEF/EPIGEN-Brazil initiative. **The Pharmacogenomics Journal**, v. 18, n. 6, p. 749–759, dez. 2018.

ROSS, J. P.; DION, P. A.; ROULEAU, G. A. Exome sequencing in genetic disease: recent advances and considerations. **F1000Research**, v. 9, p. 336, 6 maio 2020.

ROVITO, R. et al. Hallmarks of Severe COVID-19 Pathogenesis: A Pas de Deux Between Viral and Host Factors. **Frontiers in Immunology**, v. 13, p. 912336, 10 jun. 2022.

ROY, S. et al. Standards and Guidelines for Validating Next-Generation Sequencing Bioinformatics Pipelines. **The Journal of Molecular Diagnostics**, v. 20, n. 1, p. 4–27, jan. 2018.

RUIZ-STERNBERG, Á. M. et al. Genomic characterization of SARS-CoV-2 and its association with clinical outcomes: a 1-year longitudinal study of the pandemic in Colombia. **International Journal of Infectious Diseases**, v. 116, p. 91–100, mar. 2022.

SCHAAN, A.P., GUSMÃO, L., JANNUZZI, J. *et al.* New insights on intercontinental origins of paternal lineages in Northeast Brazil. *BMC Evol Biol* **20**, 15 (2020). https://doi.org/10.1186/s12862-020-1579-9

SCHMIDT, S. M. The role of iron in viral infections. **Frontiers in Bioscience**, v. 25, n. 4, p. 893–911, 2020.

SCHWARZE, K. et al. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. **Genetics in Medicine**, v. 20, n. 10, p. 1122–1130, out. 2018.

SECOLIN, R. et al. Distribution of local ancestry and evidence of adaptation in admixed populations. **Scientific Reports**, v. 9, n. 1, p. 13900, dez. 2019.

SECOLIN, R. et al. Genetic variability in COVID-19-related genes in the Brazilian population. **Human Genome Variation**, v. 8, n. 1, p. 15, dez. 2021.

SHELTON, J. F. et al. Trans-ancestry analysis reveals genetic and nongenetic associations with COVID-19 susceptibility and severity. **Nature Genetics**, v. 53, n. 6, p. 801–808, jun. 2021.

SHEN, B. et al. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. **Cell**, v. 182, n. 1, p. 59- 72.e15, jul. 2020.

SHEN, Y. et al. Ancestral origins are associated with SARS-CoV-2 susceptibility and protection in a Florida patient population. **PLOS ONE**, v. 18, n. 1, p. e0276700, 17 jan. 2023.

SHERRY, S. T. dbSNP: the NCBI database of genetic variation. **Nucleic Acids Research**, v. 29, n. 1, p. 308–311, 1 jan. 2001.

SHRIVER, M. D. et al. Skin pigmentation, biogeographical ancestry and admixture mapping. **Human Genetics**, v. 112, n. 4, p. 387–399, abr. 2003.

SUN, X. et al. Genetic and epigenetic regulation of the non-muscle myosin light chain kinase isoform by lung inflammatory factors and mechanical stress. **Clinical Science**, v. 135, n. 7, p. 963–977, 16 abr. 2021.

SYPNIEWSKI, M. et al. Gene Variants Related to Cardiovascular and Pulmonary Diseases May Correlate with Severe Outcome of COVID-19. **International Journal of Molecular Sciences**, v. 23, n. 15, p. 8696, 4 ago. 2022.

SZILÁGYI, K. L. et al. Epigenetic contribution of the myosin light chain kinase gene to the risk for acute respiratory distress syndrome. **Translational Research**, v. 180, p. 12–21, fev. 2017.

TACTACAN, C. M. et al. The pseudokinase SgK223 promotes invasion of pancreatic ductal epithelial cells through JAK1/Stat3 signaling. **Molecular Cancer**, v. 14, n. 1, p. 139, dez. 2015.

TAM, V. et al. Benefits and limitations of genome-wide association studies. **Nature Reviews Genetics**, v. 20, n. 8, p. 467–484, ago. 2019.

TAN, L. Y.; KOMARASAMY, T. V.; RMT BALASUBRAMANIAM, V. Hyperinflammatory Immune Response and COVID-19: A Double Edged Sword. **Frontiers in Immunology**, v. 12, p. 742941, 30 set. 2021.

TATE, J. G. et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. **Nucleic Acids Research**, v. 47, n. D1, p. D941–D947, 8 jan. 2019.

THE SEVERE COVID-19 GWAS GROUP. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. **New England Journal of Medicine**, v. 383, n. 16, p. 1522–1534, 15 out. 2020.

VAN DER AUWERA, G. A.; O'CONNOR, B. D. **Genomics in the Cloud**. [s.l.] O'Reilly Media, Inc., 2020.

VAN DER MADE, C. I. et al. Presence of Genetic Variants Among Young Men With Severe COVID-19. **JAMA**, v. 324, n. 7, p. 663, 18 ago. 2020.

VAN HOUT, C. V. et al. **Whole exome sequencing and characterization of coding variation in 49,960 individuals in the UK Biobank**. [s.l.] Genomics, 9 mar. 2019. Disponível em: <http://biorxiv.org/lookup/doi/10.1101/572347>. Acesso em: 1 set. 2022.

WANG, H. et al. Local Ancestry Inference in Large Pedigrees. **Scientific Reports**, v. 10, n. 1, p. 189, dez. 2020a.

WANG, J. et al. Cytokine storm and leukocyte changes in mild versus severe SARS-CoV-2 infection: Review of 3939 COVID-19 patients in China and emerging pathogenesis and therapy concepts. **Journal of Leukocyte Biology**, v. 108, n. 1, p. 17–41, 1 jul. 2020b.

WANG, P. et al. Risk factors for severe COVID-19 in middle-aged patients without comorbidities: a multicentre retrospective study. **Journal of Translational Medicine**, v. 18, n. 1, p. 461, dez. 2020c.

WANG, Y. et al. PTCH1 mutation promotes antitumor immunity and the response to immune checkpoint inhibitors in colorectal cancer patients. **Cancer Immunology, Immunotherapy**, v. 71, n. 1, p. 111–120, jan. 2022.

WHO. **WHO COVID-19 Dashboard**. Disponível em: <https://covid19.who.int/>. Acesso em: 4 fev. 2023.

WONGKITTICHOTE, P. et al. Fatal COVID -19 infection in a patient with long-chain 3-HYDROXYACYL-COA dehydrogenase deficiency: A case report. **JIMD Reports**, v. 56, n. 1, p. 40–45, nov. 2020.

XAVIER, J. et al. The ongoing COVID-19 epidemic in Minas Gerais, Brazil: insights from epidemiological data and SARS-CoV-2 whole genome sequencing. **Emerging Microbes & Infections**, v. 9, n. 1, p. 1824–1834, 1 jan. 2020.

XIE, T.-A. et al. An integrative bioinformatics analysis for identifying hub genes associated with infection of lung samples in patients infected with SARS-CoV-2. **European Journal of Medical Research**, v. 26, n. 1, p. 146, dez. 2021.

XU, L. et al. Genetic Variant of *NFIB* is Associated With the Metastasis of Osteosarcoma in Chinese Population. **Technology in Cancer Research & Treatment**, v. 18, p. 153303381987480, 1 jan. 2019.

XU, P. et al. Germline mutations in a DNA repair pathway are associated with familial colorectal cancer. **JCI Insight**, v. 6, n. 18, p. e148931, 22 set. 2021.

YAMAMOTO, H.; HIRASAWA, A. Homologous Recombination Deficiencies and Hereditary Tumors. **International Journal of Molecular Sciences**, v. 23, n. 1, p. 348, 29 dez. 2021.

YANG, J. et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. **International Journal of Infectious Diseases**, v. 94, p. 91–95, maio 2020.

YANG, S. et al. Potentially functional variants of ERAP1, PSMF1 and NCF2 in the MHC-I-related pathway predict non-small cell lung cancer survival. **Cancer Immunology, Immunotherapy**, v. 70, n. 10, p. 2819–2833, out. 2021.

YAO, Y. et al. Genome and epigenome editing identify CCR9 and SLC6A20 as target genes at the 3p21.31 locus associated with severe COVID-19. **Signal Transduction and Targeted Therapy**, v. 6, n. 1, p. 85, dez. 2021.

YILDIRIM, Z. et al. Genetic and epigenetic factors associated with increased severity of Covid-19. **Cell Biology International**, v. 45, n. 6, p. 1158–1174, jun. 2021.

YIN, Y.; WUNDERINK, R. G. MERS, SARS and other coronaviruses as causes of pneumonia: MERS, SARS and coronaviruses. **Respirology**, v. 23, n. 2, p. 130–137, fev. 2018.

ZHAO EY, JONES M, JONES SJM. Whole-Genome Sequencing in Cancer. Cold Spring Harb Perspect Med. 2019 Mar 1;9(3):a034579. doi: 10.1101/cshperspect.a034579. PMID: 29844223; PMCID: PMC6396343.

ZEBERG, H.; PÄÄBO, S. The major genetic risk factor for severe COVID-19 is inherited from Neanderthals. **Nature**, v. 587, n. 7835, p. 610–612, 26 nov. 2020.

ZHANG, Q. et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. **Science**, v. 370, n. 6515, p. eabd4570, 23 out. 2020.

ZHENG, Z. et al. Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. **Journal of Infection**, v. 81, n. 2, p. e16–e25, ago. 2020.

ZHOU, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. **Nature**, v. 579, n. 7798, p. 270–273, 12 mar. 2020.

ZHOU, Q. et al. Early-Onset Stroke and Vasculopathy Associated with Mutations in ADA2. **New England Journal of Medicine**, v. 370, n. 10, p. 911–920, 6 mar. 2014.

ANEXOS

ANEXO 1

REFERÊNCIAS RELACIONADAS COM O PAINEL DE MARCADORES DE ANCESTRALIDADE GENÔMICA DESENVOLVIDO NO LABORATÓRIO DE GENÉTICA HUMANA E MÉDICA:

1. LEAL, DIANA FEIO DA VEIGA BORGES ; SANTANA DA SILVA, MAYARA NATÁLIA ; PASTANA, L. F. ; FERNANDES, M. R. ; ATHAYDE, A. S. C. ; FERNANDES, DÉBORA CHRISTINA RICARDO OLIVEIRA ; Silva, Cleonardo Augusto ; MODESTO, ANTÔNIO A. C. ; ASSUMPCAO, P. P. ; Sidney Santos ; SANTOS, N. P. C. . Genetic Variants of MicroRNA and DROSHA Genes in Association With the Risk of Tuberculosis in the Amazon Population. Frontiers in Genetics, v. 13, p. 850068, 2022.
2. RODRIGUES, JULIANA CARLA GOMES; FERNANDES, MARIANNE RODRIGUES ; RIBEIRO-DOS-SANTOS, ANDRÉ MAURÍCIO ; DE ARAÚJO, GILDERLANIO SANTANA ; DE SOUZA, SANDRO JOSÉ ; GUERREIRO, João Farias ; RIBEIRO-DOS-SANTOS, Ândrea ; DE ASSUMPÇÃO, PAULO PIMENTEL ; SANTOS, Ney Pereira Carneiro dos ; Santos, Sidney . Pharmacogenomic Profile of Amazonian Amerindians. Journal Of Personalized Medicine, v. 12, p. 952, 2022.
3. PASTANA, LUCAS FAVACHO ; SILVA, THAYS AMÂNCIO ; GELLEN, LAURA PATRÍCIA ALBARELLO ; VIEIRA, GIOVANA MIRANDA ; DE ASSUNÇÃO, LETÍCIA ALMEIDA ; LEITÃO, LUCIANA PEREIRA COLARES ; DA SILVA, NATASHA MONTE ; COELHO, RITA DE CÁSSIA CALDERARO ; DE ALCÂNTARA, ANGÉLICA LEITE ; VINAGRE, LUI WALLACY MORIKAWA SOUZA ; RODRIGUES, JULIANA CARLA GOMES ; BORGES LEAL, DIANA FEIO DA VEIGA ; FERNANDES, MARIANNE RODRIGUES ; DE SOUZA, SANDRO JOSÉ ; KROLL, JOSÉ EDUARDO ; RIBEIRO-DOS-SANTOS, ANDRÉ MAURICIO ; BURBANO, Rommel Mario Rodríguez ; GUERREIRO, João Farias ; DE ASSUMPÇÃO, PAULO PIMENTEL ; RIBEIRO-DOS-SANTOS, ÂNDREA CAMPOS ; dos Santos, Sidney Emanuel Batista ; DOS SANTOS, NEY PEREIRA CARNEIRO . The Genomic Profile Associated with Risk of Severe Forms of COVID-19 in Amazonian Native American Populations. Journal Of Personalized Medicine, v. 12, p. 554, 2022.
4. FERNANDES, SWENY DE S. M. ; LEITÃO, LUCIANA P. C. ; COHEN-PAES, AMANDA DE N. ; GELLEN, LAURA P. A. ; PASTANA, LUCAS F. ; DE CARVALHO, DARLEN C. ; MODESTO, ANTÔNIO A. C. ; DA COSTA, ANA C. A. ; WANDERLEY, ALAYDE V. ; LIMA, CARLOS H. V. DE ; PEREIRA, ESDRAS E. B. ; FERNANDES, MARIANNE R. ; BURBANO, ROMMEL M. R. ; DE ASSUMPÇÃO, PAULO P. ; SANTOS, SIDNEY E. B. DOS ; SANTOS, NEY P. C. DOS . The Role of SLC22A1 and Genomic Ancestry on Toxicity

during Treatment in Children with Acute Lymphoblastic Leukemia of the Amazon Region. Genes, v. 13, p. 610, 2022.

5. FERNANDES, MARIANNE RODRIGUES ; SOUZA VINAGRE, LUI WALLACY MORIKAWA ; RODRIGUES, JULIANA CARLA GOMES ; WANDERLEY, ALAYDE VIEIRA ; FERNANDES, SWENY MARINHO ; GELLEN, LAURA PATRÍCIA ALBARELLO ; ALCÂNTARA, ANGÉLICA LEITE DE ; SOUSA, BEATRIZ BRILHANTE DE ; BURBANO, Rommel Mario Rodríguez ; ASSUMPÇÃO, PAULO PIMENTEL DE ; Santos, Sidney Emanuel Batista dos ; SANTOS, Ney Pereira Carneiro dos . Correlation of Genetic Variants and the Incidence, Prevalence and Mortality Rates of Acute Lymphoblastic Leukemia. Journal Of Personalized Medicine, v. 12, p. 370, 2022.
6. PEREIRA, ESDRAS E. B. ; LEITÃO, LUCIANA P. C. ; ANDRADE, ROBERTA B. ; MODESTO, ANTÔNIO A. C. ; FERNANDES, BRUNO M. ; BURBANO, ROMMEL M. R. ; ASSUMPÇÃO, PAULO P. ; FERNANDES, MARIANNE R. ; Guerreiro, João F. ; SANTOS, SIDNEY E. B. DOS ; SANTOS, NEY P. C. DOS . UGT1A1 Gene Polymorphism Contributes as a Risk Factor for Lung Cancer: A Pilot Study with Patients from the Amazon. Genes, v. 13, p. 493, 2022.
7. CEREJA PANTOJA, KARLA BEATRIZ CARDIAS ; AZEVEDO, TEREZA CRISTINA DE BRITO ; CARVALHO, DARLEN CARDOSO DE ; MONTE, NATASHA ; COHEN PAES, AMANDA DE NAZARÉ ; BARROS, MARIA CLARA DA COSTA ; VINAGRE, LUI WALLACY MORIKAWA SOUZA ; FREITAS, ANA ROSA SALES DE ; BURBANO, Rommel Mario Rodríguez ; ASSUMPÇÃO, PAULO PIMENTEL DE ; Santos, Sidney Emanuel Batista dos ; FERNANDES, MARIANNE RODRIGUES ; SANTOS, Ney Pereira Carneiro dos . Impact of Variants in the ATIC and ARID5B Genes on Therapeutic Failure with Imatinib in Patients with Chronic Myeloid Leukemia. Genes, v. 13, p. 330, 2022.
8. COHEN-PAES, AMANDA DE NAZARÉ ; DE CARVALHO, DARLEN CARDOSO ; PASTANA, LUCAS FAVACHO ; DOBBIN, ELIZABETH AYRES FRAGOSO ; MOREIRA, FABIANO CORDEIRO ; DE SOUZA, TATIANE PIEDADE ; FERNANDES, MARIANNE RODRIGUES ; LEAL, DIANA FEIO DA VEIGA BORGES ; DE SÁ, ROBERTA BORGES ANDRADE ; DE ALCÂNTARA, ANGÉLICA LEITE ; GUERREIRO, João Farias ; RIBEIRO-DOS-SANTOS, Ândrea ; dos Santos, Sidney Emanuel Batista ; DE ASSUMPÇÃO, PAULO PIMENTEL ; DOS SANTOS, NEY PEREIRA CARNEIRO . Characterization of PCLO Gene in Amazonian Native American Populations. Genes, v. 13, p. 499, 2022.
9. PORCHERA, DÉBORA C. R. F. ; LEAL, DIANA F. V. B. ; BRAGA, ANA C. O. ; PINTO, PABLO D. C. ; SANTANA DA SILVA, MAYARA N. ; BEZERRA SANTOS, LUCAS C. ; BRAGA DA SILVA, CINTIA H. ; DA COSTA, GIOVANA E. ; BARROS, MARIA CLARA DA C. ; ATHAYDE, AIDALUCY DO S. C. DE ; DE N COHEN-PAES, AMANDA ; DA SILVA, CLEONARDO A. ; DE ASSUMPÇÃO, PAULO P. ; RIBEIRO-DOS-SANTOS, ÂNDREA K. C. ; DOS SANTOS, SIDNEY E. B. ; FERNANDES, MARIANNE R. ; DOS SANTOS, NEY P. C. . Association of the rs4646994 in ACE gene with susceptibility to tuberculosis in a region of the Brazilian Amazon. TRANSLATIONAL MEDICINE COMMUNICATIONS, v. 7, p. 7-10, 2022.
10. LEITÃO, LUCIANA P. C. ; DE CARVALHO, DARLEN C. ; RODRIGUES, JULIANA C. G. ; FERNANDES, MARIANNE R. ; WANDERLEY, ALAYDE

V. ; VINAGRE, LUI W. M. S. ; DA SILVA, NATASHA M. ; PASTANA, LUCAS F. ; GELLEN, LAURA P. A. ; ASSUNÇÃO, MATHEUS C. E. ; FERNANDES, SWENY S. M. ; PEREIRA, ESDRAS E. B. ; RIBEIRO-DOS-SANTOS, ANDRÉ M. ; Guerreiro, João F. ; RIBEIRO-DOS-SANTOS, Ândrea ; DE ASSUMPÇÃO, PAULO P. ; DOS SANTOS, SIDNEY E. B. ; DOS SANTOS, NEY P. C. . Identification of Genomic Variants Associated with the Risk of Acute Lymphoblastic Leukemia in Native Americans from Brazilian Amazonia. Journal Of Personalized Medicine, v. 12, p. 856, 2022.

11. ANDRADE, ROBERTA B. ; CAVALCANTE, GIOVANNA C. ; AMADOR, MARCOS A. T. ; MOREIRA, FABIANO CORDEIRO ; Khayat, André S. ; ASSUMPÇÃO, PAULO P. ; RIBEIRO-DOS-SANTOS, Ândrea ; Santos, Ney P. C. ; Santos, Sidney . The Search for Cancer Biomarkers: Assessing the Distribution of INDEL Markers in Different Genetic Ancestries. Current Issues in Molecular Biology (Online), v. 44, p. 2275-2286, 2022.
12. RODRIGUES, JULIANA CARLA GOMES ; FERNADES, MARIANNE RODRIGUES ; BARRA, WILLIAMS FERNANDES ; PEREIRA, ESDRAS EDGAR BATISTA ; RIBEIRO-DOS-SANTOS, ANDRÉ ; ASSUMPÇÃO, PAULO PIMENTEL ; GUERREIRO, João Farias ; RIBEIRO-DOS-SANTOS, ANDREA ; Santos, Sidney Emanuel Batista dos ; SANTOS, Ney Pereira Carneiro dos . Identification of pharmacogenomic variants associated with oncology treatments in Brazilian Amazonian Amerindians.. JOURNAL OF CLINICAL ONCOLOGY, v. 40, p. e15082-e15082, 2022.
13. VIDAL, AMANDA FERREIRA ; FERRAZ, RAFAELLA SOUSA ; EL-HUSNY, ANTONETTE ; SILVA, CAIO SANTOS ; VINASCO-SANDOVAL, TATIANA ; MAGALHÃES, LEANDRO ; RAIOL-MORAES, MILENE ; BARRA, WILLIAMS FERNANDES ; PEREIRA, CYNTHIA LARA BRITO LINS ; DE ASSUMPÇÃO, PAULO PIMENTEL ; DE BRITO, LEONARDO MIRANDA ; VIALLE, RICARDO ASSUNÇÃO ; Santos, Sidney ; RIBEIRO-DOS-SANTOS, Ândrea ; RIBEIRO-DOS-SANTOS, ANDRÉ M. . Comprehensive analysis of germline mutations in northern Brazil: a panel of 16 genes for hereditary cancer-predisposing syndrome investigation. BMC CANCER, v. 21, p. 363, 2021.
14. FERNANDES, MARIANNE RODRIGUES ; RODRIGUES, JULIANA CARLA GOMES ; DOBBIN, ELIZABETH AYRES FRAGOSO ; PASTANA, LUCAS FAVACHO ; DA COSTA, DANIELLE FEIO ; BARRA, WILLIAMS FERNANDES ; MODESTO, ANTÔNIO ANDRÉ CONDE ; DE ASSUMPÇÃO, PAULA BARAÚNA ; DA COSTA SILVA, ARTUR LUIZ ; dos Santos, Sidney Emanuel Batista ; BURBANO, Rommel Mario Rodriguez ; DE ASSUMPÇÃO, PAULO PIMENTEL ; DOS SANTOS, NEY PEREIRA CARNEIRO . Influence of FPGS, ABCC4, SLC29A1, and MTHFR genes on the pharmacogenomics of fluoropyrimidines in patients with gastrointestinal cancer from the Brazilian Amazon. CANCER CHEMOTHERAPY AND PHARMACOLOGY, v. 88, p. 837-844, 2021.
15. DOBBIN, ELIZABETH AYRES FRAGOSO ; MEDEIROS, JÉSSYCA AMANDA GOMES ; COSTA, MARTA SOLANGE CAMARINHA RAMOS ; RODRIGUES, JULIANA CARLA GOMES ; GUERREIRO, João Farias ; KROLL, JOSÉ EDUARDO ; SOUZA, SANDRO JOSÉ DE ; DE ASSUMPÇÃO, PAULO PIMENTEL ; RIBEIRO-DOS-SANTOS, Ândrea ; Santos, Sidney Emanuel Batista dos ; BURBANO, Rommel Mario Rodríguez ; FERNANDES, MARIANNE RODRIGUES ; SANTOS, Ney Pereira Carneiro dos . Identification

of Variants (rs11571707, rs144848, and rs11571769) in the BRCA2 Gene Associated with Hereditary Breast Cancer in Indigenous Populations of the Brazilian Amazon. Genes, v. 12, p. 142, 2021.
16. MONTE, NATASHA ; PANTOJA, KARLA B. C. C. ; RODRIGUES, JULIANA C. G. ; CARVALHO, DARLEN C. ; AZEVEDO, TEREZA C. B. ; PEREIRA, ESDRAS E. B. ; ASSUMPÇÃO, PAULO P. ; SANTOS, SIDNEY E. B. ; FERNANDES, MARIANNE R. ; Santos, Ney P. C. . Polymorphisms in the CYP2A6 and ABCC4 genes are associated with a protective effect on chronic myeloid leukemia in the Brazilian Amazon population. MOLECULAR GENETICS & GENOMIC MEDICINE, v. 9, p. 1694, 2021.
17. FERNANDES, MARIANNE RODRIGUES ; RODRIGUES, JULIANA CARLA GOMES ; MAROÑAS, OLALLA ; LATORRE-PELLICER, ANA ; CRUZ, RAQUEL ; GUERREIRO, João Farias ; BURBANO, Rommel Mario Rodriguez ; ASSUMPÇÃO, PAULO PIMENTEL DE ; RIBEIRO-DOS-SANTOS, Ândrea ; Santos, Sidney Emanuel Batista dos ; CARRACEDO, ANGEL ; SANTOS, Ney Pereira Carneiro dos . Genetic Diversity of Drug-Related Genes in Native Americans of the Brazilian Amazon. Pharmacogenomics and Personalized Medicine, v. Volume 14, p. 117-133, 2021.
18. RODRIGUES, JULIANA CARLA GOMES ; PINTO, PABLO ; LEITÃO, LUCIANA PEREIRA COLARES ; VINAGRE, LUI WALLACY MORIKAWA SOUZA ; MONTE, NATASHA ; FERNANDES, MARIANNE RODRIGUES ; KHAYAT, André Salim ; DE ASSUMPÇÃO, PAULO PIMENTEL ; SANTOS, Ney Pereira Carneiro dos ; Santos, Sidney Emanuel Batista dos . Influence of APOE locus on poor prognosis of COVID-19. HELIYON, v. 7, p. e07379, 2021.
19. LEITÃO, LUCIANA P. C. ; SOUZA, TATIANE P. ; RODRIGUES, JULIANA C. G. ; FERNANDES, MARIANNE R. ; Santos, Sidney ; Santos, Ney P. C. . The Metabolization Profile of the CYP2D6 Gene in Amerindian Populations: A Review. Genes, v. 11, p. 262, 2020.
20. CARVALHO, DARLEN CARDOSO ; WANDERLEY, ALAYDE VIEIRA ; MELLO JUNIOR, FERNANDO AUGUSTO RODRIGUES ; SANTOS, ANDRÉ MAURICIO RIBEIRO DOS ; LEITÃO, LUCIANA PEREIRA COLARES ; SOUZA, TATIANE PIEDADE DE ; CASTRO, AMANDA DE NAZARÉ COHEN LIMA DE ; MAGALHÃES, LEANDRO LOPES DE ; FERNANDES, MARIANNE RODRIGUES ; CARVALHO JUNIOR, JOÃO AUGUSTO NUNES DE ; KHAYAT, André Salim ; Santos, Sidney ; ASSUMPÇÃO, PAULO PIMENTEL DE ; DOS SANTOS, NEY PEREIRA CARNEIRO . Association of genes ARID5B, CEBPE and folate pathway with acute lymphoblastic leukemia in a population from the Brazilian Amazon region. LEUKEMIA RESEARCH REPORTS, v. 13, p. 100188, 2020.
21. CARDOSO DE CARVALHO, DARLEN ; PEREIRA COLARES LEITÃO, LUCIANA ; MELLO JUNIOR, FERNANDO AUGUSTO RODRIGUES ; VIEIRA WANDERLEY, ALAYDE ; SOUZA, TATIANE PIEDADE DE ; BORGES ANDRADE DE SÁ, ROBERTA ; COHEN-PAES, AMANDA ; RODRIGUES FERNANDES, MARIANNE ; Santos, Sidney ; SALIM KHAYAT, ANDRÉ ; PIMENTEL DE ASSUMPÇÃO, PAULO ; PEREIRA CARNEIRO DOS SANTOS, NEY . Association between the TPMT*3C (rs1142345) Polymorphism and the Risk of Death in the Treatment of Acute Lymphoblastic Leukemia in Children from the Brazilian Amazon Region. Genes, v. 11, p. 1132, 2020.

22. REIS, LARISSA B. ; Bakos, Renato M. ; VIANNA, FERNANDA S. L. ; MACEDO, GABRIEL S. ; JACOVAS, VANESSA C. ; RIBEIRO-DOS-SANTOS, ANDRÉ M. ; Santos, Sidney ; BAKOS, LÚCIO ; Ashton-Prolla, Patricia . Skin pigmentation polymorphisms associated with increased risk of melanoma in a case-control sample from southern Brazil. BMC CANCER, v. 20, p. 1069, 2020.
23. DE CARVALHO, DARLEN CARDOSO ; WANDERLEY, ALAYDE VIEIRA ; DOS SANTOS, ANDRÉ MAURICIO RIBEIRO ; MOREIRA, FABIANO CORDEIRO ; DE SÁ, ROBERTA BORGES ANDRADE ; FERNANDES, MARIANNE RODRIGUES ; MODESTO, ANTONIO ANDRÉ CONDE ; DE SOUZA, TATIANE PIEDADE ; COHEN-PAES, AMANDA ; LEITÃO, LUCIANA PEREIRA COLARES ; RODRIGUES, JULIANA CARLA GOMES ; DA SILVA, ARTUR LUIZ DA COSTA ; GUERREIRO, João Farias ; Santos, Sidney ; KHAYAT, André Salim ; DE ASSUMPÇÃO, PAULO PIMENTEL ; DOS SANTOS, NEY PEREIRA CARNEIRO . Characterization of pharmacogenetic markers related to Acute Lymphoblastic Leukemia toxicity in Amazonian native Americans population. Scientific Reports, v. 10, p. 10292, 2020.
24. KHAYAT, André Salim ; DE ASSUMPÇÃO, PAULO PIMENTEL ; MEIRELES KHAYAT, BRUNA CLAUDIA ; THOMAZ ARAÚJO, TAÍSSA MAÍRA ; BATISTA-GOMES, JÉSSICA ALMEIDA ; IMBIRIBA, LUCIANA CARVALHO ; ISHAK, GERALDO ; DE ASSUMPÇÃO, PAULA BARAÚNA ; MOREIRA, FABIANO CORDEIRO ; BURBANO, ROMMEL RODRIGUEZ ; RIBEIRO-DOS-SANTOS, ANDRÉ ; RIBEIRO-DOS-SANTOS, ÂNDREA KELLY ; DOS SANTOS, NEY PEREIRA CARNEIRO ; DOS SANTOS, SIDNEY EMMANUEL BATISTA . ACE2 polymorphisms as potential players in COVID-19 outcome. PLoS One, v. 15, p. e0243887, 2020.
25. RIBEIRO-DOS-SANTOS, ANDRÉ M. ; VIDAL, AMANDA FERREIRA ; VINASCO-SANDOVAL, TATIANA ; GUERREIRO, JOÃO ; Santos, Sidney ; RIBEIRO-DOS-SANTOS, Ândrea ; DE SOUZA, SANDRO J. . Exome Sequencing of Native Populations From the Amazon Reveals Patterns on the Peopling of South America. Frontiers in Genetics, v. 11, p. 548537, 2020.
26. RODRIGUES, JULIANA CARLA GOMES ; SOUZA, TATIANE PIEDADE DE ; PASTANA, LUCAS FAVACHO ; RIBEIRO DOS SANTOS, ANDRÉ MAURÍCIO ; FERNANDES, MARIANNE RODRIGUES ; PINTO, PABLO ; WANDERLEY, ALAYDE VIEIRA ; SOUZA, SANDRO JOSÉ DE ; KROLL, JOSÉ EDUARDO ; PEREIRA, ADENILSON LEÃO ; MAGALHÃES, LEANDRO ; MERCÊS, LAÍS REIS DAS ; VIDAL, AMANDA FERREIRA ; VINASCO-SANDOVAL, TATIANA ; CAVALCANTE, GIOVANNA CHAVES ; GUERREIRO, João Farias ; ASSUMPÇÃO, PAULO PIMENTEL DE ; RIBEIRO-DOS-SANTOS, Ândrea ; Santos, Sidney ; SANTOS, Ney Pereira Carneiro dos . Identification of NUDT15 gene variants in Amazonian Amerindians and admixed individuals from northern Brazil. PLoS One, v. 15, p. e0231651, 2020.
27. LEAL, DIANA FEIO DA VEIGA BORGES ; SANTANA DA SILVA, MAYARA NATÁLIA ; FERNANDES, DÉBORA CRISTINA RICARDO DE OLIVEIRA ; RODRIGUES, JULIANA CARLA GOMES ; BARROS, MARIA CLARA DA COSTA ; PINTO, PABLO DIEGO DO CARMO ; PASTANA, LUCAS FAVACHO ; SILVA, CLEONARDO AUGUSTO DA ; FERNANDES, MARIANNE RODRIGUES ; ASSUMPÇÃO, PAULO PIMENTEL DE ; Santos,

Sidney Emanuel Batista dos ; DOS SANTOS, NEY PEREIRA CARNEIRO . Amerindian genetic ancestry as a risk factor for tuberculosis in an amazonian population. PLoS One, v. 15, p. e0236033, 2020.
28. CAVALCANTE, GIOVANNA CHAVES ; DE MORAES, MILENE RAIOL ; VALENTE, Cristina Maria Duarte ; SILVA, CAIO SANTOS ; MODESTO, ANTÔNIO ANDRÉ CONDE ; DE ASSUMPÇÃO, PAULA BARAÚNA ; DE ASSUMPÇÃO, PAULO PIMENTEL ; Santos, Sidney ; RIBEIRO-DOS-SANTOS, Ândrea . Investigation of INDEL variants in apoptosis: the relevance to gastric cancer. BMC Medical Genetics, v. 21, p. 207, 2020.
29. CASTRO, A. N. C. L. ; FERNANDES, M. R. ; CARVALHO, D. C. ; SOUZA, T. P. ; RODRIGUES, J. C. G. ; DE SÁ, ROBERTA BORGES ANDRADE ; MODESTO, A. A. C. ; Sidney Santos ; ASSUMPCAO, P. P. ; SANTOS, N. P. C. . Polymorphisms of xenobiotic-metabolizing and transporter genes, and the risk of gastric and colorectal cancer in an admixed population from the Brazilian Amazon. American Journal of Translational Research, v. 12, p. 6626-6636, 2020.
30. FIGUEIREDO, EDUARDO ANTÔNIO ; LOYOLA, LEONOR CASILLA ; BELANGERO, PAULO SANTORO ; RIBEIRO'DOS'SANTOS, ÂNDREA KELY CAMPOS ; Santos, Sidney Emanuel Batista ; COHEN, CARINA ; WAJNSZTEJN, ANDRE ; MARTINS DE OLIVEIRA, ADRIELLE ; SMITH, MARÍLIA CARDOSO ; CASTRO POCHINI, ALBERTO ; ANDREOLI, CARLOS VICENTE ; EJNISMAN, BENNO ; COHEN, MOISES ; LEAL, MARIANA FERREIRA . Rotator cuff tear susceptibility is associated with variants in genes involved in tendon extracellular matrix homeostasis. JOURNAL OF ORTHOPAEDIC RESEARCH, v. 38, p. 192-201, 2019.
31. RODRIGUES, JULIANA CARLA GOMES ; FERNANDES, MARIANNE RODRIGUES ; GUERREIRO, João Farias ; DA SILVA, ARTUR LUIZ DA COSTA ; RIBEIRO-DOS-SANTOS, Ândrea ; Santos, Sidney ; SANTOS, Ney Pereira Carneiro dos . Polymorphisms of ADME-related genes and their implications for drug safety and efficacy in Amazonian Amerindians. Scientific Reports, v. 9, p. 7201, 2019.
32. ASTUR, DIEGO COSTA ; ANDRADE, EDILSON ; ARLIANI, GUSTAVO GONÇALVES ; DEBIEUX, PEDRO ; LOYOLA, LEONOR CASILLA ; dos Santos, Sidney Emanuel Batista ; BURBANO, Rommel Mario Rodriguez ; LEAL, MARIANA FERREIRA ; COHEN, MOISES . The potential European genetic predisposition for non-contact anterior cruciate ligament injury. KNEE SURGERY SPORTS TRAUMATOLOGY ARTHROSCOPY, v. 1, p. 1, 2018.
33. GUIMARÃES, LILIAN O. ; FERNANDES, FRANCISCO ; MONTEIRO, ELIANA F. ; CURADO, IZILDA ; HOLCMAN, MARCIA M. ; WUNDERLICH, GERHARD ; Santos, Sidney E. ; SOLER, JÚLIA M. ; KIRCHGATTER, KARIN . Influence of polymorphisms in toll-like receptors (TLRs) on malaria susceptibility in low-endemic area of the Atlantic Forest, São Paulo, Brazil. ACTA TROPICA, v. 182, p. 309-316, 2018.
34. ANDRADE, ROBERTA B. ; AMADOR, MARCOS A. T. ; CAVALCANTE, GIOVANNA C. ; LEITÃO, LUCIANA P. C. ; FERNANDES, MARIANNE R. ; MODESTO, ANTÔNIO A. C. ; MOREIRA, FABIANO C. ; Khayat, André S. ; ASSUMPÇÃO, PAULO P. ; RIBEIRO-DOS-SANTOS, Ândrea ; Santos, Sidney ; Santos, Ney P. C. . Estimating Asian Contribution to the Brazilian Population: A New Application of a Validated Set of 61 Ancestry Informative Markers. G3-Genes Genomes Genetics, v. 8, p. 3577-3582, 2018.

35. DE CARVALHO, DARLEN CARDOSO ; WANDERLEY, ALAYDE VIEIRA ; DOS SANTOS, ANDRÉ MAURICIO RIBEIRO ; FERNANDES, MARIANNE RODRIGUES ; COHEN LIMA DE CASTRO, AMANDA DE NAZARÉ ; LEITÃO, LUCIANA PEREIRA COLARES ; DE CARVALHO, JOÃO AUGUSTO NUNES ; DE SOUZA, TATIANE PIEDADE ; KHAYAT, André Salim ; dos Santos, Sidney Emanuel Batista ; DE ASSUMPÇÃO, PAULO PIMENTEL ; DOS SANTOS, NEY PEREIRA CARNEIRO . Pharmacogenomics and variations in the risk of toxicity during the consolidation/maintenance phases of the treatment of pediatric B-cell leukemia patients from an admixed population in the Brazilian Amazon. LEUKEMIA RESEARCH, v. 74, p. 10-13, 2018.
36. COHEN, CARINA ; LEAL, MARIANA FERREIRA ; LOYOLA, LEONOR CASILLA ; Santos, Sidney Emanuel Batista ; RIBEIRO-DOS-SANTOS, Ândrea Kely Campos ; BELANGERO, PAULO SANTORO ; FIGUEIREDO, EDUARDO ANTÔNIO ; WAJNSZTEJN, ANDRE ; DE OLIVEIRA, ADRIELLE MARTINS ; SMITH, MARÍLIA CARDOSO ; ANDREOLI, CARLOS VICENTE ; DE CASTRO POCHINI, ALBERTO ; COHEN, MOISES ; EJNISMAN, BENNO ; FALOPPA, FLÁVIO . Genetic variants involved in extracellular matrix homeostasis play a role in the susceptibility to frozen shoulder: a case-control study. JOURNAL OF ORTHOPAEDIC RESEARCH, v. 1, p. 1, 2018.
37. ARRIFANO, GABRIELA P. F. ; MARTÍN-DOIMEADIOS, ROSA C. R. ; JIMÉNEZ-MORENO, MARÍA ; FERNÁNDEZ-TRUJILLO, SERGIO ; AUGUSTO-OLIVEIRA, MARCUS ; SOUZA-MONTEIRO, JOSÉ R. ; MACCHI, BARBARELLA M. ; ALVAREZ-LEITE, JACQUELINE I. ; DO NASCIMENTO, JOSÉ L. M. ; AMADOR, MARCOS T. ; Santos, Sidney ; RIBEIRO-DOS-SANTOS, Ândrea ; SILVA-PEREIRA, LIZ C. ; ORIÁ, REINALDO B. ; CRESPO-LOPEZ, MARIA E. . Genetic Susceptibility to Neurodegeneration in Amazon: Apolipoprotein E Genotyping in Vulnerable Populations Exposed to Mercury. Frontiers in Genetics, v. 9, p. 1, 2018.
38. D?ELIA, MARIA PAULA BARBIERI ; BRANDÃO, MARCELA CALIXTO ; DE ANDRADE RAMOS, BRUNA RIBEIRO ; DA SILVA, MÁRCIA GUIMARÃES ; MIOT, LUCIANE DONIDA BARTOLI ; dos Santos, Sidney Emanuel Batista ; MIOT, HÉLIO AMANTE . African ancestry is associated with facial melasma in women: a cross-sectional study. BMC Medical Genetics, v. 18, p. 1-7, 2017.
39. LOPES, C. B. ; BARROSO, R. F. F. ; BURBANO, Rommel Mario Rodríguez ; GARCIA, P. A. ; PINTO, P. D. C. ; SANTOS, N. P. C. ; SANTOS, S E B ; RIBEIRO-DOS-SANTOS, A. K. C. . Effect of ancestry on interleukin-10 haplotypes in chronic periodontitis. Frontiers in Bioscience (Elite Edition), v. 1, p. 276-285, 2017.
40. MARQUES, D. ; COSTA, L. R. F. ; COSTA, L. L. F. ; CORREA, R. S. ; BORGES, A. M. P. ; ITO, F. R. ; RAMOS, C. C. O. ; BORTOLIN, R. H. ; LUCHESSI, A. D. ; RIBEIRO-DOS-SANTOS, A. K. C. ; Sidney E. dos Santos ; SILBERGER, V. N. . Association of insertion-deletions polymorphisms with colorectal cancer risk and clinical features. WORLD JOURNAL OF GASTROENTEROLOGY, v. 23, p. 6854, 2017.
41. SORTICA, VINICIUS A ; LINDENAU, JULIANA D ; CUNHA, MARISTELA G ; O OHNISHI, MARIA DEISE ; R VENTURA, ANA MARIA ; RIBEIRO-DOS-SANTOS, Ândrea KC ; Santos, Sidney EB ; GUIMARÃES, LUCIANO SP ; HUTZ, MARA H . SLCO1A2, SLCO1B1 and SLCO2B1 polymorphisms

influences chloroquine and primaquine treatment in Plasmodium vivax malaria. PHARMACOGENOMICS, v. 10, p. 1, 2017.
42. MASOTTI, C. ; BRITO, L.A. ; NICA, A.C. ; LUDWIG, K.U. ; NUNES, K. ; SAVASTANO, C.P. ; MALCHER, C. ; FERREIRA, S.G. ; KOBAYASHI, G.S. ; BUENO, D.F. ; ALONSO, N. ; FRANCO, D. ; ROJAS-MARTINEZ, A. ; DOS SANTOS, S.E. ; GALANTE, P.A. ; MEYER, D. ; HÜNEMEIER, T. ; MANGOLD, E. ; DERMITZAKIS, E.T. ; PASSOS-BUENO, M.R. . MRPL53 , a New Candidate Gene for Orofacial Clefting, Identified Using an eQTL Approach. JOURNAL OF DENTAL RESEARCH, v. 11, p. 002203451773580, 2017.
43. DA SILVA, ELLEN MORENO ; FERNANDES, MARIANNE RODRIGUES ; DE CARVALHO, DARLEN CARDOSO ; LEITAO, LUCIANA PEREIRA COLARES ; CAVALCANTE, GIOVANNA CHAVES ; PEREIRA, ESDRAS EDGAR BATISTA ; MODESTO, ANTÔNIO ANDRÉ CONDE ; GUERREIRO, João Farias ; DE ASSUMPÇÃO, PAULO PIMENTEL ; dos Santos, Sidney Emanuel Batista ; DOS SANTOS, NEY PEREIRA CARNEIRO . Effect of genetic ancestry to the risk of susceptibility to gastric cancer in a mixed population of the Brazilian Amazon. BMC RESEARCH NOTES, v. 10, p. 1, 2017.
44. CAVALCANTE, GIOVANNA C ; AMADOR, MARCOS AT ; SANTOS, ANDRÉ M RIBEIRO DOS ; CARVALHO, DARLEN C ; ANDRADE, ROBERTA B ; PEREIRA, ESDRAS EB ; FERNANDES, MARIANNE R ; COSTA, DANIELLE F ; SANTOS, NEY PC ; ASSUMPÇÃO, PAULO P ; SANTOS, ÂNDREA RIBEIRO DOS ; Santos, Sidney . Analysis of 12 variants in the development of gastric and colorectal cancers. WORLD JOURNAL OF GASTROENTEROLOGY, v. 23, p. 8533-8543, 2017.
45. DA SILVA, ELLEN MORENO ; FERNANDES, MARIANNE RODRIGUES ; DE CARVALHO, DARLEN CARDOSO ; LEITAO, LUCIANA PEREIRA COLARES ; CAVALCANTE, GIOVANNA CHAVES ; PEREIRA, ESDRAS EDGAR BATISTA ; MODESTO, ANTÔNIO ANDRÉ CONDE ; GUERREIRO, João Farias ; DE ASSUMPÇÃO, PAULO PIMENTEL ; dos Santos, Sidney Emanuel Batista ; DOS SANTOS, NEY PEREIRA CARNEIRO . Correction to: Effect of genetic ancestry to the risk of susceptibility to gastric cancer in a mixed population of the Brazilian Amazon. BMC RESEARCH NOTES, v. 10, p. 29187240, 2017.
46. GUIMARÃES, LILIAN O. ; BAJAY, MIKLOS MAXIMILIANO ; MONTEIRO, ELIANA F. ; WUNDERLICH, GERHARD ; Santos, Sidney E. ; KIRCHGATTER, KARIN . Genetic ancestry effects on the distribution of toll-like receptors (TLRs) gene polymorphisms in a population of the Atlantic Forest, São Paulo, Brazil. HUMAN IMMUNOLOGY, v. 79, p. 101-108, 2017.
47. RAMOS, BRUNA RIBEIRO DE ANDRADE ; MENDES, NIELE DIAS ; TANIKAWA, ALINE AKI ; AMADOR, MARCOS ANTÔNIO TRINDADE ; SANTOS, Ney Pereira Carneiro dos ; Santos, Sidney Emanuel Batista dos ; CASTELLI, ERICK C. ; WITKIN, STEVEN S. ; SILVA, MÁRCIA GUIMARÃES DA . Ancestry informative markers and selected single nucleotide polymorphisms in immunoregulatory genes on preterm labor and preterm premature rupture of membranes: a case control study. BMC Pregnancy and Childbirth (Online), v. 16, p. PMC4743423, 2016.
48. AMADOR, MARCOS A. T. ; CAVALCANTE, GIOVANNA C. ; Santos, Ney P. C. ; Gusmão, Leonor ; Guerreiro, João F. ; RIBEIRO-DOS-SANTOS, Ândrea ; Santos, Sidney . Distribution of allelic and genotypic frequencies of IL1A, IL4,

NFKB1 and PAR1 variants in Native American, African, European and Brazilian populations. BMC Research Notes, v. 9, p. 101, 2016.
49. RAMOS, BRUNA RIBEIRO DE ANDRADE ; D?ELIA, MARIA PAULA BARBIERI ; AMADOR, MARCOS ANTÔNIO TRINDADE ; SANTOS, Ney Pereira Carneiro ; Santos, Sidney Emanuel Batista ; DA CRUZ CASTELLI, ERICK ; WITKIN, STEVEN S. ; MIOT, HÉLIO AMANTE ; MIOT, LUCIANE DONIDA BARTOLI ; DA SILVA, MÁRCIA GUIMARÃES . Neither self-reported ethnicity nor declared family origin are reliable indicators of genomic ancestry. Genetica ('s-Gravenhage), v. 3, p. 001, 2016.
50. DE SEIXAS SANTOS NASTRI, ANA CATHARINA ; DE MELLO MALTA, FERNANDA ; DINIZ, MÁRCIO AUGUSTO ; YOSHINO, ALESSANDRA ; ABE-SANDES, KIYOKO ; dos Santos, Sidney Emanuel Batista ; DE CASTRO LYRA, ANDRÉ ; CARRILHO, FLAIR JOSÉ ; PINHO, JOÃO RENATO REBELLO . Association of IFNL3 and IFNL4 polymorphisms with hepatitis C virus infection in a population from southeastern Brazil. Archives of Virology, v. 3, p. 001, 2016.
51. CASSIANO, GUSTAVO C. ; FURINI, ADRIANA A. C. ; CAPOBIANCO, MARCELA P. ; STORTI-MELO, LUCIANE M. ; CUNHA, MARISTELA G. ; KANO, FLORA S. ; CARVALHO, LUZIA H. ; SOARES, IRENE S. ; Santos, Sidney E. ; PÓVOA, MARINETE M. ; MACHADO, RICARDO L. D. . Polymorphisms in B Cell Co-Stimulatory Genes Are Associated with IgG Antibody Responses against Blood-Stage Proteins of Plasmodium vivax. Plos One, v. 11, p. e0149581, 2016.
52. MAESTRI, ALVINO ; SORTICA, VINICIUS ALBUQUERQUE ; FERREIRA, DEIMY LIMA ; DE ALMEIDA FERREIRA, JESSYLENE ; AMADOR, MARCOS ANTÔNIO TRINDADE ; DE MELLO, WYLLER ALENCAR ; Santos, Sidney Emanuel Batista ; SOUSA, RITA CATARINA MEDEIROS . The His131Arg substitution in the FCGR2A gene (rs1801274) is not associated with the severity of influenza A(H1N1)pdm09 infection. BMC Research Notes, v. 9, p. 9:296, 2016.
53. WAGNER, SANDRINE C. ; LINDENAU, JULIANA D. ; CASTRO, SIMONE M. DE ; SANTIN, ANA PAULA ; ZALESKI, CARINA F. ; AZEVEDO, LAURA A. ; RIBEIRO DOS SANTOS, ÂNDREA K. C. ; DOS SANTOS, SIDNEY E.B. ; Hutz, Mara H. . High Frequency of Hb E-Saskatoon (HBB : c.67G->-A) in Brazilians: A New Genetic Origin?. Hemoglobin, v. 2, p. 1-3, 2016.
54. EL-HUSNY, ANTONETTE ; RAIOL-MORAES, MILENE ; AMADOR, MARCOS ; RIBEIRO-DOS-SANTOS, ANDRÉ M. ; MONTAGNINI, ANDRÉ ; BARBOSA, SILVANIRA ; Silva, Artur ; Assumpção, Paulo ; ISHAK, GERALDO ; Santos, Sidney ; PINTO, PABLO ; CRUZ, ALINE ; RIBEIRO-DOS-SANTOS, Ândrea . CDH1 mutations in gastric cancer patients from northern Brazil identified by Next- Generation Sequencing (NGS). Genetics and Molecular Biology (online version), v. 2, p. 189-198, 2016.
55. ESCOBAR, GABRIELA F. ; ARRAES, JOSE AROLDO A. ; BAKOS, Lucio ; Ashton-Prolla, Patricia ; Giugliani, Roberto ; CALLEGARI-JACQUES, SIDIA MARIA ; Santos, Sidney ; Bakos, Renato M. . Polymorphisms in CYP19A1 and NFKB1 genes are associated with cutaneous melanoma risk in southern Brazilian patients. MELANOMA RESEARCH, v. 26, p. 348-353, 2016.
56. VIEIRA-MACHADO, CAMILLA D. ; DE CARVALHO, FLAVIA M. ; SANTANA DA SILVA, LUIZ C. ; Dos Santos, Sidney E. ; MARTINS, CLAUDIA ; POLETTA, FERNANDO A. ; MEREB, JUAN C. ; VIEIRA,

ALEXANDRE R. ; CASTILLA, EDUARDO E. ; ORIOLI, IÊDA M. . Analysis of the genetic ancestry of patients with oral clefts from South American admixed populations. European Journal of Oral Sciences, v. 00, p. 6, 2016.
57. FURINI, ADRIANA ANTÔNIA DA CRUZ ; CASSIANO, GUSTAVO CAPATTI ; PETROLINI CAPOBIANCO, MARCELA ; dos Santos, Sidney Emanuel Batista ; DANTAS MACHADO, RICARDO LUIZ . Frequency of TNFA , INFG , and IL10 Gene Polymorphisms and Their Association with Malaria Vivax and Genomic Ancestry. MEDIATORS OF INFLAMMATION, v. 2016, p. 1-12, 2016.
58. SORTICA, VINICIUS A ; LINDENAU, JULIANA D ; CUNHA, MARISTELA G ; OHNISHI, MARIA DO ; VENTURA, Ana Maria R ; RIBEIRO-DOS-SANTOS, Ândrea KC ; Santos, Sidney EB ; GUIMARÃES, LUCIANO SP ; HUTZ, MARA H . The effect of SNPs in CYP450 in chloroquine/primaquine Plasmodium vivax malaria treatment. PHARMACOGENOMICS, v. 17, p. 17, 2016.
59. TOMAZ, FRANCIELE MAIRA MOREIRA BATISTA ; DA CRUZ FURINI, ADRIANA ANTÔNIA ; CAPOBIANCO, MARCELA PETROLINI ; PÓVOA, Marinete Marins ; TRINDADE, PAMELLA CRISTINA ALVES ; FRAGA, VALÉRIA DALTIBARI ; CONCEIÇÃO, LUCIANA MORAN ; DE AZEVEDO, LUCAS RIBEIRO ; OLIANI, SÔNIA MARIA ; CASSIANO, GUSTAVO CAPATTI ; CAVASINI, CARLOS EUGÊNIO ; dos Santos, Sidney Emanuel Batista ; MACHADO, RICARDO LUIZ DANTAS . Humoral immune responses against the malaria vaccine candidate antigen Plasmodium vivax AMA-1 and IL-4 gene polymorphisms in individuals living in an endemic area of the Brazilian Amazon. Cytokine, v. 4, p. S1043-4666, 2015.
60. CARDENA, M M S G ; RIBEIRO-DOS-SANTOS, A K ; SANTOS, S E B ; MANSUR, A J ; BERNARDEZ-PEREIRA, S ; SANTOS, P C J L ; PEREIRA, A C ; FRIDMAN, C . Mitochondrial and genomic ancestry are associated with etiology of heart failure in Brazilian patients. Journal of Human Hypertension, v. 5, p. 1, 2015.
61. VIEIRA, P. ; BURBANO, Rommel Mario Rodriguez ; FERNANDES, DÉBORA CHRISTINA RICARDO OLIVEIRA ; MONTENEGRO, R. C. ; Santos, Sidney ; SORTICA, VINICIUS A ; Assumpção, Paulo ; RIBEIRO-DOS-SANTOS, Ândrea ; CARVALHO, A. ; SANTOS, N. P. C. . Population stratification effect on cancer susceptibility in an admixed population from brazilian Amazon.. Anticancer Research, v. 35, p. 2009, 2015.
62. FERREIRA PALHA, TERESINHA DE JESUS BRABO ; RIBEIRO RODRIGUES, ELZEMAR MARTINS ; CAVALCANTE, GIOVANNA CHAVES ; MARRERO, ANDREA ; SOUZA, ILÍADA RAINHA DE ; SEKI UEHARA, CLINEU JULIEN ; SILVEIRA DA MOTTA, CARLOS HENRIQUE ARES ; KOSHIKENE, DANIELA ; SILVA, DAYSE APARECIDA DA ; CARVALHO, ELIZEU FAGUNDES DE ; CHEMALE, GUSTAVO ; FREITAS, JORGE M. ; ALEXANDRE, LÍDIA ; PARANAIBA, RENATO T.F. ; SOLER, MIRELLA PERRUCCIO ; Santos, Sidney . Population genetic analysis of insertion-deletion polymorphisms in a Brazilian population using the Investigator DIPplex kit.. Forensic Science International. Genetics (Print), v. 4, p. 1, 2015.
63. CARVALHO, DARLEN C. ; WANDERLEY, ALAYDE V. ; AMADOR, MARCOS A.T. ; FERNANDES, MARIANNE R. ; CAVALCANTE, GIOVANNA C. ; PANTOJA, KARLA B.C.C. ; MELLO, FERNANDO A.R. ; DE ASSUMPÇÃO, PAULO P. ; Khayat, André S. ; RIBEIRO-DOS-SANTOS,

Ândrea ; Santos, Sidney ; DOS SANTOS, NEY P.C. . Amerindian genetic ancestry and INDEL polymorphisms associated with susceptibility of childhood B-cell eukemia in an admixed population from the Brazilian Amazon. Leukemia Research, v. 08, p. LR5442, 2015.

64. MAESTRI, ALVINO ; DOS SANTOS, MIRLEIDE CORDEIRO ; RIBEIRO-RODRIGUES, ELZEMAR M ; DE MELLO, WYLLER ALENCAR ; SOUSA, RITA CATARINA MEDEIROS ; DOS SANTOS, SIDNEY EMANUEL ; SORTICA, VINICIUS ALBUQUERQUE . The CCR5-32 (rs333) polymorphism is not a predisposing factor for severe pandemic influenza in the Brazilian admixed population. BMC Research Notes, v. 8, p. 326, 2015.
65. SOUZA, ELISA S.T. DE ; ARAÚJO, LUIZA F. DE ; ALENCAR, DAYSE O. DE ; SANTOS, SIDNEY E.B. DOS ; SILVA JR, WILSON A. ; FERREIRA, CRISTIANE A. ; BADDINI-MARTINEZ, JOSÉ . Does ethnic ancestry play a role in smoking?. Anais da Academia Brasileira de Ciências (Impresso), v. 00, p. 00-00, 2015.
66. FERNANDES, DÉBORA CHRISTINA RICARDO OLIVEIRA ; SANTOS, Ney Pereira Carneiro ; MORAES, MILENE RAIOL ; BRAGA, ANA CRISTINA OLIVEIRA ; Silva, Cleonardo Augusto ; RIBEIRO-DOS-SANTOS, ANDREA ; Santos, Sidney . Association of the CYP2B6 gene with anti-tuberculosis drug-induced hepatotoxicity in a Brazilian Amazon population. International Journal of Infectious Diseases, v. 33, p. 28-31, 2015.
67. DE ARAUJO, TÂNIA KAWASAKI ; SECOLIN, RODRIGO ; FÉLIX, TÊMIS MARIA ; DE SOUZA ; BARROS FONTES, MARSHALL ÍTALO ; MONLLEÓ, ISABELLA LOPES ; DE SOUZA, JOSIANE ; FETT-CONTE, AGNES CRISTINA ; RIBEIRO, ERLANE MARQUES ; XAVIER, ANA CAROLINA ; DE REZENDE, ADRIANA AUGUSTO ; SIMIONI, MILENA ; RIBEIRO-DOS-SANTOS, Ândrea Kely Campos ; dos Santos, Sidney Emanuel Batista ; GIL-DA-SILVA-LOPES, VERA LÚCIA . A multicentric association study between 39 genes and nonsyndromic cleft lip and palate in a Brazilian population. Journal of Cranio-Maxillo-Facial Surgery, v. 08, p. 001, 2015.
68. PINTO, P. D. ; SALGADO, CG ; SANTOS, N. P. C. ; Santos, Sidney ; RIBEIRO-DOS-SANTOS, A. . Influence of Genetic Ancestry on INDEL Markers of NFKβ1, CASP8, PAR1, IL4 and CYP19A1 Genes in Leprosy Patients. PLoS Neglected Tropical Diseases (Online), v. 9, p. e0004050, 2015.
69. SILVA, AYLLA N. L. M. ; CARDOSO, GREICE L. ; CUNHA, DANIELE A. ; DINIZ, ISABELA G. ; Santos, Sidney E.B. ; ANDRADE, GABRIELA B. ; TRINDADE, SAIDE M.S. ; CARDOSO, MARIA DO SOCORRO O. ; FRANCÊS, LARISSA T.V.M. ; Guerreiro, João F. . The Spectrum of -Thalassemia Mutations in a Population from the Brazilian Amazon. Hemoglobin, v. sep, p. 1-5, 2015.
70. CASSIANO, GUSTAVO CAPATTI ; SANTOS, EDUARDO JOSÉ MELO ; MAIA, MARIA HELENA THOMAZ ; DA CRUZ FURINI, ADRIANA ; STORTI-MELO, LUCIANE MORENO ; TOMAZ, FRANCIELE MAIRA BATISTA ; TRINDADE, PAMELLA CRISTINA ALVES ; CAPOBIANCO, MARCELA PETROLINI ; AMADOR, MARCOS ANTÔNIO TRINDADE ; VIANA, GISELLE MARIA RACHID ; PÓVOA, Marinete Marins ; Santos, Sidney Emanuel Batista ; MACHADO, RICARDO LUIZ DANTAS . Impact of population admixture on the distribution of immune response co-stimulatory genes polymorphisms in a Brazilian population. Human Immunology, v. 76, p. 737, 2015.

71. MAESTRI, ALVINO ; SORTICA, VINICIUS ALBUQUERQUE ; TOVO-RODRIGUES, LUCIANA ; SANTOS, MIRLEIDE CORDEIRO ; BARBAGELATA, LUANA ; MORAES, MILENE RAIOL ; ALENCAR DE MELLO, WYLLER ; Gusmão, Leonor ; SOUSA, RITA CATARINA MEDEIROS ; EMANUEL BATISTA DOS SANTOS, SIDNEY ; TANG, JULIAN W. . Siaα2-3Galβ1- Receptor Genetic Variants Are Associated with Influenza A(H1N1)pdm09 Severity. Plos One, v. 10, p. e0139681, 2015.
72. PASKULIN, DIEGO DAVILA ; GIACOMAZZI, JULIANA ; ACHATZ, MARIA ISABEL ; COSTA, SANDRA ; REIS, RUI MANOEL ; HAINAUT, PIERRE ; dos Santos, Sidney Emanuel Batista ; Ashton-Prolla, Patricia . Ancestry of the Brazilian TP53 c.1010G>A (p.Arg337His, R337H) Founder Mutation: Clues from Haplotyping of Short Tandem Repeats on Chromosome 17p. Plos One, v. 10, p. e0143262, 2015.
73. CAVALCANTE, GIOVANNA CHAVES ; FREITAS, Natalle Do Socorro da Costa ; RIBEIRO-DOS-SANTOS, A. M. ; CARVALHO, DARLEN C. ; SILVA, E. M. ; ASSUMPÇÃO, PAULO PIMENTEL ; SANTOS, Andrea Kely Campos Ribeiro dos ; SANTOS, SIDNEY E. B. ; SANTOS, N. P. C. . INVESTIGATION OF POTENTIALLY DELETERIOUS ALLELES FOR RESPONSE TO CANCER TREATMENT WITH 5-FLUOROURACIL. Anticancer Research, v. 12, p. 6871, 2015.
74. DE SOUZA, LILIANE TODESCHINI ; KOWALSKI, THAYNE WOYCINCK ; FERRARI, JÉSSICA ; MONLLEO, ISABELLA LOPES ; RIBEIRO, ERLANE MARQUES ; DE SOUZA, JOSIANE ; FETT-CONTE, AGNES CRISTINA ; DE ARAUJO, TÂNIA KAWASAKI ; GIL DA SILVA-LOPES, VERA ; RIBEIRO-DOS-SANTOS, Ândrea Kely Campos ; Santos, Sidney Emanuel Batista dos ; FÉLIX, TÊMIS MARIA . Study of IRF6 and 8q24 region in non-syndromic oral clefts in the Brazilian population. Oral Diseases, v. 22, p. n/a-n/a, 2015.
75. OTA, VANESSA K. ; BELLUCCO, FERNANDA T. ; GADELHA, ARY ; SANTORO, MARCOS L. ; NOTO, CRISTIANO ; CHRISTOFOLINI, DENISE M. ; ASSUNÇÃO, IDAIANE B. ; YAMADA, KAREN M. ; RIBEIRO-DOS-SANTOS, ÂNDREA K. ; Santos, Sidney ; MARI, JAIR J. ; SMITH, MARÍLIA A. C. ; MELARAGNO, MARIA I. ; BRESSAN, RODRIGO A. ; SATO, JOÃO R. ; JACKOWSKI, ANDREA P. ; BELANGERO, SINTIA I. . PRODH Polymorphisms, Cortical Volumes and Thickness in Schizophrenia. Plos One, v. 9, p. e87686, 2014.
76. GOMES, LARISSA LUZ ; MOREIRA, FABIANO CORDEIRO ; Hamoy, Igor Guerreiro ; Santos, Sidney ; Assumpção, Paulo ; SANTANA, ÁDAMO L. ; SANTOS, ÂNDREA . Identification of miRNAs Expression Profile in Gastric Cancer Using Self-Organizing Maps (SOM). Bioinformation print, v. 10, p. 246-250, 2014.
77. COLARES, VINÍCIUS SARDÃO ; TITAN, SILVIA MARIA DE OLIVEIRA ; PEREIRA, ALEXANDRE DA COSTA ; MALAFRONTE, PATRÍCIA ; CARDENA, MARI M. ; Santos, Sidney ; SANTOS, PAULO C. ; FRIDMAN, CÍNTIA ; BARROS, RUI TOLEDO ; WORONIK, VIKTÓRIA . MYH9 and APOL1 Gene Polymorphisms and the Risk of CKD in Patients with Lupus Nephritis from an Admixture Population. Plos One, v. 9, p. e87716, 2014.
78. CARDENA, MARI M.S.G. ; RIBEIRO-DOS-SANTOS, Ândrea ; Santos, Sidney ; MANSUR, ALFREDO J. ; PEREIRA, ALEXANDRE C. ; FRIDMAN, CINTIA . Amerindian genetic ancestry is associated with higher survival rates compared

to African and European ancestry in Brazilian patients with heart failure. International Journal of Cardiology (Print), v. 176, p. 527-528, 2014.
79. CARDOSO, Greice Lemos ; DINIZ, ISABELA GUERREIRO ; MARTINS DA SILVA, AYLLA NÚBIA LIMA ; CUNHA, DANIELE ALMEIDA ; DA SILVA JUNIOR, JOSIVALDO SOARES ; CARVALHO UCHÔA, CAMILA TAVARES ; dos Santos, Sidney Emanuel Batista ; TRINDADE, SAIDE MARIA SARMENTO ; CARDOSO, MARIA DO SOCORRO DE OLIVEIRA ; GUERREIRO, João Farias . DNA polymorphisms at BCL11A, HBS1L-MYB and Xmn1-HBG2 site loci associated with fetal hemoglobin levels in sickle cell anemia patients from Northern Brazil. Blood Cells, Molecules & Diseases (Print), v. 53, p. 176-179, 2014.
80. DE ALENCAR, DAYSE OLIVEIRA ; NETTO, CRISTINA ; Ashton-Prolla, Patricia ; Giugliani, Roberto ; RIBEIRO-DOS-SANTOS, Ândrea ; PEREIRA, FERNANDA ; MATTE, URSULA ; SANTOS, NEY ; Santos, Sidney . Fabry disease: Evidence for a regional founder effect of the GLA gene mutation 30delG in Brazilian patients. Molecular Genetics and Metabolism Reports, v. 1, p. 414-421, 2014.
81. OLAVARRIA, VIVIANA ; MELO DOS SANTOS, EDUARDO ; BATISTA DOS SANTOS, SIDNEY ; GRASSI, FERNANDA ; KRUCHEWESKY, RAMON ; DE MORAES SOBRINO PÔRTO, LUÍS ; GAO, XIAOJAN ; CARRINGTON, MARY ; GALVÃO-CASTRO, BERNARDO . Do African and European ancestry and polymorphism of HLA class I play an important role in controlling HTLV-1 proviral load in admixed cohorts from Salvador, Brazil?. Retrovirology, v. 11, p. P76, 2014.
82. OTA, VANESSA KIYOMI ; BELLUCCO, F. T. ; CHRISTOFOLINI, D. ; MARI, J. ; BRESSAN, R. ; RIBEIRO-DOS-SANTOS, A. ; Santos, Sidney ; BREEN, G. ; BELANGERO, S. I. N. ; SMITH, MARILIA DE ARRUDA CARDOSO . Candidate genes for schizophrenia in a mixed Brazilian population using pooled DNA. Psychiatry Research (Print), p. 0165-1781, 2013.
83. BRASIL-COSTA, I. ; ALENCAR, Dayse de Oliveira ; MORAES, Milene Raiol de ; PESSOA, I. A. ; BRITO, A. W. M. ; JATI, S. R. ; Santos, Sidney ; BURBANO, Rommel Mario Rodriguez ; RIBEIRO-DOS-SANTOS, A. K. C. . Molecular Characterization of TP53 Gene in Human Populations Exposed to Low-Dose Ionizing Radiation. BioMed Research International, v. 2013, p. 1-9, 2013.
84. VASCONCELOS, JANAINA MOTA DE ; MÓIA, LIZOMAR DE JESUS MAUÉS PEREIRA ; AMARAL, IVANETE DO SOCORRO ABRAÇADO ; MIRANDA, ESTHER CASTELLO BRANCO MELLO ; CICALISETAKESHITA, LOUISE YUKARI ; OLIVEIRA, LAYANNA FREITAS DE ; MENDES, LILIAN DE ARAÚJO MELO ; SASTRE, DANUTA ; TAMEGÃO-LOPES, BRUNA PEDROSO ; PEDROZA, LARYSSE SANTA ROSA DE AQUINO ; Santos, Sidney Emanuel Batista dos ; SOARES, MANOEL DO CARMO PEREIRA ; ARAÚJO, MARIALVA TEREZA FERREIRA DE ; BANDEIRA, CAMILA LUCAS ; SILVA, ADRIANA MARIA PAIXÃO DE SOUSA DA ; MEDEIROS, ZILENE LAMEIRA DE ; SENA, LEONARDO ; DEMACHKI, Samia ; SANTOS, Eduardo José Melo dos . Association of killer cell immunoglobulin-like receptor polymorphisms with chronic hepatitis C and responses to therapy in Brazil. Genetics and Molecular Biology (Impresso), v. 36, p. 022-027, 2013.

85. CARDENA, MARI M. S. G. ; RIBEIRO-DOS-SANTOS, Ândrea ; Santos, Sidney ; MANSUR, ALFREDO J. ; PEREIRA, ALEXANDRE C. ; FRIDMAN, CINTIA . Assessment of the Relationship between Self-Declared Ethnicity, Mitochondrial Haplogroups and Genomic Ancestry in Brazilian Individuals. Plos One, v. 8, p. e62005, 2013.
86. LOPES, T. R. ; Santos, Sidney ; Resque, Rafael Lima ; RIBEIRO-DOS-SANTOS, A. ; PINTOA, G. R. ; YOSHIOKA, France Keiko Nascimento . Population data of the 46 insertion-deletion (INDEL) loci in population in Piauí State, Northeastern Brazil. Forensic Science International. Genetics (Print), v. 1, p. 2013.07.009, 2013.
87. FERNANDES, M. R. ; CARVALHO, D. C. ; SANTOS, Ândrea Kely Campos Ribeiro dos ; Santos, Sidney ; ASSUMPCAO, P. ; BURBANO, Rommel Mario Rodriguez ; SANTOS, N. P. C. . Association of Slow Acetylation Profile of NAT2 with Breast and Gastric Cancer Risk in Brazil. Anticancer Research, v. 33, p. 3683/1791-7530-3690, 2013.
88. SANTOS, Ney Pereira Carneiro dos ; Callegari-Jacques, S. M. C. ; RIBEIRO-DOS-SANTOS, A. ; Silva, Cleonardo Augusto ; VALLINOTO, Antônio Carlos Rosário ; Fernandes ; CARVALHO, D. C. ; Santos, Sidney ; HUTZ, Mara Helena . N-acetyl transferase 2 and cytochrome P450 2E1 genes and isoniazid-induced hepatotoxicity in Brazilian patients. The International Journal of Tuberculosis and Lung Disease, v. 17, p. 499-504, 2013.
89. GARCIA, P. ; ALENCAR, D. ; PINTO, P. ; SANTOS, N. ; SALGADO, C. ; SORTICA, V. A. ; HUTZ, M. H. ; RIBEIRO DOS SANTOS, A. ; SANTOS, S. . Haplotypes of IL10 gene (A-1082G, C-819T and C-592A) as potential protection factors in leprosy patients.. Clinical and Vaccine Immunology, v. 20, p. 00334-13, 2013.
90. SORTICA, VINICIUS A. ; CUNHA, MARISTELA G. ; OHNISHI, MARIA D.O. ; SOUZA, JOSE M. ; Ribeiro-dos-Santos, Ândrea K.C. ; Santos, Sidney E.B. ; Hutz, Mara H. . Role of IL6, IL12B and VDR gene polymorphisms in Plasmodium vivax malaria severity, parasitemia and gametocytemia levels in an Amazonian Brazilian population. Cytokine, v. 1, p. 2013.09.014, 2013.
91. RIBEIRO-DOS-SANTOS, A. M. ; SOUSA, J. E. ; ALMEIDA, R. ; ALENCAR, D. O. ; Barbosa, Maria Silvanira ; GUSMÃO, L. ; SILVA JR, Wilson A ; SOUZA, S. J. ; Silva, Artur ; RIBEIRO-DOS-SANTOS, A. ; Darnet, Sylvain ; Santos, Sidney . High-Throughput Sequencing of a South American Amerindian. Plos One, v. 8, p. e83340, 2013.
92. PICANCO, J. B. ; RAIMANN, P. E. ; Paskulin GA ; ALVAREZ, L. ; AMORIM, A. ; SANTOS, S. ; ALHO, C. S. . Tri-allelic pattern at the TPOX locus: A familial study. Gene (Amsterdam), p. 535-538, 2013.
93. FRIEDRICH, DEISE C. ; SANTOS, SIDNEY E. B. ; RIBEIRO-DOS-SANTOS, Ã?NDREA K. C. ; Hutz, Mara H. . Correction: Several Different Lactase Persistence Associated Alleles and High Diversity of the Lactase Gene in the Admixed Brazilian Population. Plos One, v. 8, p. eCollection 20, 2013.
94. Francez, Pablo Abdon da Costa ; Ribeiro-Rodrigues, Elzemar Martins ; dos Santos, Sidney Emanuel Batista . Allelic frequencies and statistical data obtained from 48 AIM INDEL loci in an admixed population from the Brazilian Amazon. Forensic Science International. Genetics (Print), v. 6, p. 132-135, 2012.
95. Batista-dos-Santos, Sergio ; MORAES, Milene Raiol de ; Santos, Sidney ; CUNHA, Maristela Gomes da ; RIBEIRO-DOS-SANTOS, Ândrea Kely Campos

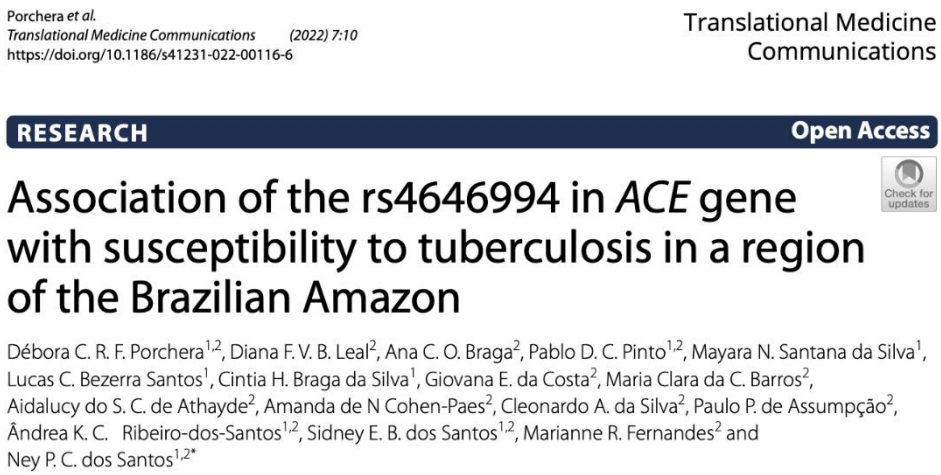
. Real-time PCR diagnosis of Plasmodium vivax among blood donors. Malaria Journal (Online), v. 11, p. 345-358, 2012.
96. PINTO, P. D. ; SALGADO, Claudio Guedes ; SANTOS, Ney Pereira Carneiro dos ; Alencar, Dayse O. ; Santos, Sidney ; HUTZ, M. ; RIBEIRO-DOS-SANTOS, A. . Polymorphisms in the CYP2E1 and GSTM1 Genes as Possible Protection Factors for Leprosy Patients. Plos One, v. 7, p. e47498, 2012.
97. OTA, VANESSA KIYOMI ; SPÍNDOLA, LETÍCIA NERY ; GADELHA, ARY ; SANTOS FILHO, AIRTON FERREIRA DOS ; SANTORO, MARCOS LEITE ; CHRISTOFOLINI, DENISE MARIA ; BELLUCCO, FERNANDA TEIXEIRA ; RIBEIRO-DOS-SANTOS, ÂNDREA KELY ; Santos, Sidney ; MARI, JAIR DE JESUS ; MELARAGNO, MARIA ISABEL ; BRESSAN, RODRIGO AFFONSECA ; SMITH, MARILIA DE ARRUDA CARDOSO ; BELANGERO, SINTIA IOLE . DRD1 rs4532 polymorphism: A potential pharmacogenomic marker for treatment response to antipsychotic drugs. Schizophrenia Research (Print), v. 06, p. xxx-xxx, 2012.
98. FRIEDRICH, DEISE C. ; SANTOS, SIDNEY E. B. ; RIBEIRO-DOS-SANTOS, ÂNDREA K. C. ; Hutz, Mara H. . Several Different Lactase Persistence Associated Alleles and High Diversity of the Lactase Gene in the Admixed Brazilian Population. Plos One, v. 7, p. e46520, 2012.
99. KIMURA, LILIAN ; Ribeiro-Rodrigues, Elzemar Martins ; DE MELLO AURICCHIO, MARIA TERESA BALESTER ; VICENTE, JOÃO PEDRO ; Batista Santos, Sidney Emanuel ; MINGRONI-NETTO, REGINA CÉLIA . Genomic ancestry of rural African-derived populations from Southeastern Brazil. American Journal of Human Biology, v. 11, p. n/a-n/a, 2012.
100. SORTICA, VINICIUS A ; CUNHA, MARISTELA G ; OHNISHI, MARIA DEISE ; SOUZA, JOSE M ; RIBEIRO-DOS-SANTOS, Ândrea KC ; SANTOS, NEY PC ; CALLEGARI-JACQUES, SÍDIA M ; Santos, Sidney EB ; HUTZ, MARA H . IL1B, IL4R, IL12RB1 and TNF gene polymorphisms are associated with plasmodium vivax malaria in Brazil. Malaria Journal (Online), v. 11, p. 409, 2012.
101. Francez, Pablo Abdon da Costa ; RODRIGUES, Elzemar Martins Ribeiro ; Frazão, Gleycianne Furtado ; Borges, Nathalia Danielly dos Reis ; Santos, Sidney . Allelic frequencies and statistical data obtained from 12 codis STR loci in an admixed population of the Brazilian Amazon. Genetics and Molecular Biology (Impresso), v. 34, p. 35-39, 2011.
102. Pedroza, L. ; Sauma, M. ; Vasconcelos, J. ; Takeshita, L. ; Ribeiro-Rodrigues, E. ; Sastre, D. ; Barbosa, C. ; Chies, J. ; Veit, T. ; Lima, C. ; Oliveira, L. ; Henderson, B. ; Castro, A. ; Maia, M. ; Barbosa, F. ; Santos, S. ; Guerreiro, J. ; Sena, L. ; SANTOS, E. ; Santos, Sidney E.B. . Systemic lupus erythematosus: Association with KIR and SLC11A1 polymorphisms, ethnic predisposition and influence in clinical manifestations at onset revealed by ancestry genetic markers in an urban Brazilian population. Lupus (Basingstoke), v. 20, p. 265-273, 2011.
103. MACIEL, Luana Gomes Lopes ; Ribeiro-Rodrigues, Elzemar ; Santos, Ney P. C. ; RIBEIRO-DOS-SANTOS, Ândrea ; GUERREIRO, Joao Farias ; Santos, Sidney . Afro-Derived Amazonian Populations: Inferring Continental Ancestry and Population Substructure. Human Biology (Print), v. 83, p. 627-636, 2011.
104. Bakos, Renato M. ; Besch, Robert ; Zoratto, Gabriela G. ; Godinho, Janaína M. ; Mazzotti, Nicolle G. ; Ruzicka, Thomas ; BAKOS, Lucio ; Santos, Sidney E. ; Ashton-Prolla, Patricia ; Berking, Carola ; Giugliani, Roberto . The CDKN2A

p.A148T variant is associated with cutaneous melanoma in Southern Brazil. Experimental Dermatology, v. 20, p. 890-893, 2011.
105. Sintia Iole ; F.T.S ; D.M ; M.I. ; M.A.C. ; ARY ; VANESSA KIYOMI ; RIBEIRO-DOS-SANTOS, Ândrea KC ; SANTOS, S. E. B. ; R. . The UFD1L rs5992403 polymorphism is associated with age at onset of schizophrenia. Journal of Psychiatric Research, p. 01, 2010.
106. Santos, Ney P.C. ; Ribeiro-Rodrigues, Elzemar M. ; Ribeiro-dos-Santos, Ândrea K.C. ; Pereira, Rui ; Gusmão, Leonor ; Amorim, António ; Guerreiro, Joáo F. ; Zago, Marco A. ; Matte, Cecília ; Hutz, Mara H. ; Santos, Sidney E.B. . Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INDEL) ancestry-informative marker (AIM) panel. Human Mutation, v. 31, p. 184-190, 2010.

ANEXO 2

REFERÊNCIAS BIBLIOGRÁFICAS RELACIONADAS COM MINHA FORMAÇÃO (2021 a 2023):

1. Artigo



Porchera *et al.*
Translational Medicine Communications (2022) 7:10
https://doi.org/10.1186/s41231-022-00116-6

Translational Medicine Communications

RESEARCH Open Access

Association of the rs4646994 in *ACE* gene with susceptibility to tuberculosis in a region of the Brazilian Amazon

Débora C. R. F. Porchera[1,2], Diana F. V. B. Leal[2], Ana C. O. Braga[2], Pablo D. C. Pinto[1,2], Mayara N. Santana da Silva[1], Lucas C. Bezerra Santos[1], Cintia H. Braga da Silva[1], Giovana E. da Costa[2], Maria Clara da C. Barros[2], Aidalucy do S. C. de Athayde[2], Amanda de N Cohen-Paes[2], Cleonardo A. da Silva[2], Paulo P. de Assumpção[2], Ândrea K. C. Ribeiro-dos-Santos[1,2], Sidney E. B. dos Santos[1,2], Marianne R. Fernandes[2] and Ney P. C. dos Santos[1,2*]

2. Resumos publicados em anais de congressos

    1. BARROS, M. C; GOMES, D.H. ; PINHO, C. T. ; PINHEIRO, J.A.S. ; SILVA, C. H. B. ; SILVA, C.S. ; VIDAL, A. F. ; SOUZA, S.J. ; SOUSA, J.E.S. ; SANTOS, A. K. C. R. . EXOME SEQUENCING IN SAMPLES OF COVID-19 PATIENTS. In: 67th Brazilian Congress of Genetics, 2022, Natal. 67th Brazilian Congress of Genetics, 2022.

    2. SILVA, C.S. ; PINTO, P. D. ; SILVA, C. H. B. ; BARROS, M. C ; PINHO, C. T. ; MAGALHAES, L. L. ; ALBUQUERQUE, R. S. ; VIEIRA, J. R. S. ;

SANTOS, A. K. C. R. ; SOUSA, M. S. . Saliva sample for SARS-CoV-2 detection using RT-PCR. In: 66th Brazilian Congress of Genetics, 2021, Online. 66th Brazilian Congress of Genetics, 2021.

3. PINHO, C. T. ; BARROS, M. C ; SILVA, C.S. ; CLOSSET, L. ; MAGALHAES, L. L. ; SILVA, C. H. B. ; ANDRADE, V. G. D. ; MONTEIRO, A. C. N. ; SANTOS, A. K. C. R. . Analysis of SARS-CoV-2 strains dispersion in the Amazon. In: 66th Brazilian Congress of Genetics, 2021, Online. 66th Brazilian Congress of Genetics, 2021.

4. BARROS, M. C; FERRAZ, R. S. ; PINHO, C. T. ; MAGALHAES, L. L. ; SILVA, C. H. B. ; SILVA, C.S. ; SALGADO, C. G. ; MEDEIROS, R. C. ; SANTOS, A. K. C. R. . IDENTIFICATION OF SARS-COV-2 LINEAGES ASSOCIATED WITH SEVERE CLINICAL CONDITIONS IN SAMPLES OF COVID-19 PATIENTS. In: 66th Brazilian Congress of Genetics, 2021, Online. 66th Brazilian Congress of Genetics, 2021.

5. LEAL, DIANA FEIO DA VEIGA BORGES ; SILVA, M.N.S. ; FERNANDES, M. R. ; VINAGRE, L. W. M. S. ; BARROS, M. C ; PASTANA, L. F. ; SANTOS, N. P. C. . Genetic association of polymorphisms with the risk of tuberculosis in Brazilian Amazon population. In: 66th Brazilian Congress of Genetics, 2021, Online. 66th Brazilian Congress of Genetics, 2021.