



Universidad del
Rosario

Escuela de Medicina y
Ciencias de la Salud



UNIVERSIDAD CES

Un compromiso con la excelencia

Factores Genéticos y no genéticos asociados a la severidad clínica de la infección por SARS-CoV-2 y COVID-19 prolongado en una muestra de la población colombiana.

**Universidad del Rosario
Escuela de medicina y ciencias de la salud**

**Universidad CES
Facultad de medicina**

Maestría en Epidemiología

Bogotá D.C, Junio 2023



Universidad del
Rosario

Escuela de Medicina y
Ciencias de la Salud



UNIVERSIDAD CES

Un compromiso con la excelencia

Factores Genéticos y no genéticos asociados a la severidad clínica de la infección por SARS-CoV-2 en una muestra de la población colombiana.

**Trabajo de investigación para optar al título de
Máster en epidemiología**

Presentado por

Mariana Angulo Aguado
mariana.anguloa@urosario.edu.co

Juan Camilo Carrillo Martínez
Juanc.carrillo@urosario.edu.co

Tutores temáticos
Dora Janeth Fonseca
dora.fonseca@urosario.edu.co

Oscar Javier Ortega
oortegar@unal.edu.co

Tutor metodológico

María Cristina Ospina
mariacr.ospina@urosario.edu.co

Universidad del Rosario
Escuela de medicina y ciencias de la salud

Universidad CES
Facultad de medicina

Maestría en epidemiología
Bogotá D.C, Diciembre 2022

La Universidad del Rosario y la Universidad CES no se hacen responsables de los conceptos emitidos por los investigadores en su trabajo, sólo velará por el rigor científico, metodológico y ético en aras de la búsqueda de la verdad y la justicia.

CONTENIDO

RESUMEN.....	6
1.FORMULACIÓN DEL PROBLEMA.....	7
1.1 Planteamiento del problema.....	7
2.JUSTIFICACIÓN.....	9
2.1 Justificación general.....	9
2.2 Justificación específica.....	10
3. PREGUNTA DE INVESTIGACIÓN.....	10
3.1 Pregunta de investigación principal	10
3.2 Preguntas de investigación secundarias.....	10
4. MARCO TEÓRICO.....	11
4.1 Contexto de la pandemia originada por SARS-CoV-2.....	11
4.2 Epidemiología.....	12
4.3 Fisiopatología de la infección por SARS-CoV-2.....	13
4.4 Caracterización clínica de la enfermedad.....	15
4.5 COVID-19 prolongado.....	16
4.6. Factores clínicos	18
4.7 Factores virales y moleculares.....	19
4.8 Factores Genéticos del huésped.....	20
4.9 Estudios genéticos en Colombia.....	21
4.10 Herramientas de priorización en Salud pública y Modelos de regresión aplicados a COVID19.....	22
5. HIPÓTESIS.....	23
5.1 Hipótesis nula.....	23
5.2 Hipótesis alternativa.....	23
6. OBJETIVOS.....	23
6.1 Objetivo general.....	23
6.2 Objetivos específicos.....	23
7. METODOLOGÍA.....	24
7.1 Enfoque metodológico.....	24
7.2 Tipo de estudio.....	24
7.3 Población.....	24
7.4 Descripción de la población de estudio.....	24
7.5 Definición de la unidad de análisis.....	25
7.6 Criterios de inclusión y exclusión.....	25

7.61 Criterios de inclusión generales.....	25
7.62 Criterios de inclusión grupo casos.....	25
7.63 Criterios de inclusión grupo controles.....	26
7.6.4 Criterios de exclusión.....	26
7.7 Tamaño de la muestra.....	26
7.7.1 Método de muestreo.....	26
7.7.2 Cálculo del tamaño de la muestra.....	27
7.8 Definición de las variables.....	27
7.9 Técnicas de recolección de la información.....	40
7.9.1 Fuente de información.....	40
7.10 Control de sesgos.....	41
8. PLAN DE ANÁLISIS.....	42
8.1 Técnicas de procesamiento y análisis de datos.....	42
8.1.1 Análisis genético.....	42
8.1.2 Análisis estadístico.....	43
8.1.3 Análisis univariado.....	44
8.1.4 Análisis bivariado.....	44
8.1.5 Análisis multivariado.....	45
9. CONSIDERACIONES ÉTICAS.....	45
9.1 Calificación del riesgo del estudio.....	45
9.2 Método de obtención del consentimiento informado.....	45
9.3 Métodos de minimización de riesgos del sujeto.....	46
10. RESULTADOS.....	46
10.1 Definición de la muestra de estudio.....	46
10.2 Caracterización demográfica y clínica.....	47
10.3 Análisis de variables clínicas y demográficas.....	50
10.4 Análisis de calidad de los datos genéticos.....	51
10.5 Análisis de polimorfismo de nucleótido simple (SNP).....	52
10.5.1 Parámetros genético poblacionales.....	52
10.5.2 Análisis bivariado con severidad de COVID-19.....	57
10.5.3 Desequilibrio de ligamiento.....	64
10.5.4 Análisis bivariado con presencia de COVID-19 prolongado.....	66
10.6 Análisis de variantes moleculares en genes candidato.....	71
10.7 Modelos predictivos de COVID-19.....	74
10.7.1 Modelo predictivo de severidad.....	76
10.7.2 Modelo predictivo de COVID-19 prolongado.....	77
11. DISCUSIÓN.....	77
12. LIMITACIONES.....	82
13. CONCLUSIONES.....	82
14. BIBLIOGRAFÍA.....	83

TABLAS:

Tabla 0: Definición de variables.....	27
Tabla 1: Características demográficas y clínicas de la población de estudio.....	48
Tabla 2: Síntomas de COVID-19 en la población de estudio.....	48
Tabla 3: Signos y síntomas de COVID-19 en la población de estudio.....	49
Tabla 4: Características clínicas y demográficas de los pacientes con y sin COVID-19 pronogado.....	51
Tabla 5: Síntomas de COVID-19 en pacientes con y sin COVID-19 prolongado.....	51
Tabla 6: Frecuencias alélicas y genotípicas de casos y controles de severidad de COVID-19.....	53
Tabla 7: Frecuencias alélicas y genotípicas de pacientes con y sin COVID-19 prolongado.....	56
Tabla 8: Análisis bivariado entre los polimorfismos genéticos con la severidad de COVID-19.....	57
Tabla 9: Análisis bivariado entre los polimorfismos genéticos con la presencia de COVID-19 prolongado.....	66
Tabla 10: Comparación de las variantes LoF y missense patogénicas entre casos y controles por gen.....	72
Tabla 11: Comparación de los diferentes modelos predictivos clínico, genético y mixto para severidad de COVID-19.....	75
Tabla 12: Comparación de los diferentes modelos predictivos clínico, genético y mixto para presencia de COVID-19 prolongado.....	75

FIGURAS:

Figura 1: Diagrama de flujo de los participantes en el estudio.....	47
Figura 2: Análisis de desequilibrio de ligamiento SNPs Chr 3.....	65
Figura 3: Análisis de desequilibrio de ligamiento SNPs Chr17.....	65
Figura 4: Comparación curvas ROC.....	76

ANEXOS:

Tabla suplementaria 1: Totalidad de variantes genéticas identificadas en el panel de genes	
Tabla suplementaria 2: Calidad de secuenciación por genes y por pacientes	
Tabla suplementaria 3: Análisis de calidad profundidad de genes	
Tabla suplementaria 4: Totalidad de variantes genéticas identificadas en el panel de genes	

RESUMEN

Introducción: Existe una heterogénea respuesta a la infección por SARS-CoV-2 dada por variabilidad interindividual mediada por factores clínicos, virales y genéticos del huésped. Pocos estudios a nivel latinoamericano han explorado la implicación de variantes genéticas en la variabilidad fenotípica de la evolución y desenlaces de esta enfermedad. El principal objetivo del estudio fue evaluar la asociación de factores clínicos y genéticos con la severidad de COVID-19 y la presencia de COVID-19 prolongado.

Metodología: Utilizando un diseño metodológico tipo casos y controles (n = 112; 1:1). Se recolectaron variables clínicas, sociodemográficas, y genéticas a través de encuestas, la historia clínica y secuenciación de siguiente generación (NGS). Se determinó la asociación de las variables clínicas y de 81 polimorfismos de nucleótido simple (SNP) con severidad de COVID-19 y presencia de COVID-19 prolongado. Las variables significativas se usaron para construir un modelo predictivo de severidad de COVID-19 y COVID-19 prolongado.

Resultados y discusión: En concordancia con reportes previos, variables clínicas como sexo masculino, obesidad, diabetes mellitus, tos, disnea, entre otras, exhibieron asociación estadísticamente significativa con la severidad a COVID-19. Varios SNP presentaron asociación significativa con severidad, tales como rs11385942 ($p < 0.01$; OR=10.88; 95% CI=1.36-86.51), rs10490770 ($p < 0.01$; OR=9.699; 95% CI=1.20-77.89), rs35081325 ($p < 0.01$; OR=9.699; 95% CI=1.20-77.89) y rs73064425 ($p < 0.01$; OR=9.699; 95% CI=1.20-77.89). Múltiples síntomas respiratorios y sistémicos presentes durante la infección tuvieron asociación con la presencia de COVID-19 prolongado, al igual que algunos SNP como rs8178521 del gen *IL10RB* ($p < 0.01$; OR=2.51; 95% CI=1.27-4.94). Se ilustró que los modelos mixtos que integran variables clínicas y genéticas tienen el mejor poder predictivo.

1. FORMULACIÓN DEL PROBLEMA

1.1 Planteamiento del problema:

La pandemia originada en marzo 11 del 2020 por el SARS-CoV-2 (1) generó un problema de salud pública que afectó a más de 640 millones de personas y produjo la muerte de cerca de 6.5 millones de ellas alrededor del mundo (datos a noviembre de 2022) (2). En Colombia, el seguimiento de la pandemia liderada por el Instituto Nacional de Salud reveló que 6.3 millones de personas han sido confirmadas como positivas para la infección con una tasa de mortalidad de 2.2% (141,862 fallecidos). Se ha determinado diferencias en la tasa de positividad según el departamento del territorio nacional, y de manera puntual en Bogotá se presenta el mayor número de casos positivos reportados (mas de 1.8 millones de casos a noviembre de 2022) (3).

El SARS-CoV-2 es un coronavirus constituido por una hebra monocatenaria de ácido ribonucleico (ARN) en sentido positivo con una longitud de 29,891 pares de bases nucleotídicas nombrado y clasificado por el comité internacional de taxonomía de virus (ICTV) como un betacoronavirus del grupo 2B dada la similitud de su genoma en un 79.5% con SARS-CoV (4). De manera similar a otros coronavirus utiliza el receptor ACE2 como blanco de entrada y requiere de la proteólisis de la proteína S para ingresar a la célula a través de la ruta endocítica (5). Se conoce que algunas proteasas del huésped como la proteasa transmembrana serina proteasa 2 (TMPRSS2), la cathepsina L y la furina tienen un rol importante en la escisión de dicha proteína. El receptor ACE2 y TMPRSS2 se co-expresan en las células epiteliales de la vía aérea superior lo cual explica el tropismo tisular del virus y el compromiso predominante de estos tejidos (6).

Los síntomas característicos del COVID-19 incluyen fiebre, tos, fatiga, disnea, odinofagia, anosmia, ageusia, cefalea, mialgias y/o artralgias, náuseas y/o vómito, rinorrea, diarrea (7). De manera importante, se ha descrito una amplia variabilidad interindividual en el fenotipo clínico de la enfermedad que oscila desde casos asintomáticos a casos críticos e incluso fatales. Se estima que el 80% de los casos presenta enfermedad leve, el 13.1% severa y el 6.1% evolucionan a enfermedad crítica (8). Adicionalmente la presentación clínica prolongada de COVID-19, fenómeno conocido como COVID-19 prolongado, en inglés "*long-COVID*", se caracteriza por una amplia gama de síntomas que aparecen en el 30% de los pacientes con COVID-19 después de la recuperación (9) (10).

Existen varios factores implicados en esta heterogeneidad clínica, incluidos aspectos sociodemográficos, antecedentes clínicos, variantes genéticas del virus y la constitución genética del huésped (11). Dentro de los factores de riesgo que confieren mayor morbimortalidad se encuentran la edad mayor de 60 años, sexo masculino, antecedentes médicos como hipertensión arterial, diabetes mellitus, enfermedad cardiovascular, enfermedad pulmonar obstructiva crónica, cáncer, depresión y

obesidad (12) (6). Otros factores se relacionan con variantes genéticas del virus, puesto que mutaciones en la secuencia pueden originar cambios en las proteínas virales y modular el ciclo viral, virulencia, interacciones virus-células o la inmunogenicidad (13).

De manera importante, el rol de los factores genéticos del huésped en la susceptibilidad, severidad, evolución y desenlace de la enfermedad causada por SARS-CoV-2 ha sido estudiado en modelos celulares, animales y poblaciones humanas desde el surgimiento de la pandemia (11). Entre estos estudios cabe destacar la identificación de diferentes polimorfismos de nucleótido simple (SNP) en genes relacionados con la fisiopatología de la enfermedad, tanto en el ingreso del virus a la célula (*ACE2*, *TMPRSS2*, *MX1*) como con la respuesta inmunológica y el proceso inflamatorio (*TLR3*, *IRF3*, *IRF7*, *IFNAR1*, *IFNAR2*, *KLRC2*, *DPP9*, *TYK2*, *LZTFL1*) (14). Estas observaciones han llevado a concluir que la genética del huésped constituye un factor importante en el desarrollo y evolución de la enfermedad.

Para el desarrollo de COVID-19 severo/crítico se ha estimado una heredabilidad basada en polimorfismos nucleótido simple (SNP) comunes de 6.5%, siendo esta la proporción de la varianza en el fenotipo clínico de la enfermedad explicada por la variación genotípica del individuo (15). La identificación de SNP asociados con la susceptibilidad y severidad de la enfermedad se ha llevado a cabo mediante múltiples estudios de casos y controles o de asociación de genoma amplio (GWAS), donde se ha demostrado que la presencia de variantes alélicas específicas como c.8790 G>A del gen *ACE2* (rs2285666), la Inserción/Delección de la secuencia Alu del gen *ACE* (rs4646994), c.42T>C en el gen *IFITM3* (rs12252), c.388T>C en el gen *APOE* (rs429358), entre otras, están asociadas a mayor severidad de COVID-19. A la fecha se han descrito más de 1000 genes de riesgo con variantes moleculares, las cuales se estima contribuyen al 77% de la heredabilidad basada en SNP para enfermedad severa (16) (17).

Adicionalmente, estos estudios han permitido la identificación de 13 loci significativos independientes, correspondientes a regiones genómicas implicadas en desarrollo y severidad de la enfermedad. Entre ellos se incluyen el locus 9q34.2, donde se encuentran los genes codificantes para el grupo sanguíneo ABO (18) y los loci 19p13.2, 6p21.1, 19p13.3 y 3p21.31 (19). Por otra parte, en cuanto COVID-19 prolongado, la presencia de secuelas y el desenlace clínico posterior a la infección también puede estar intrínsecamente relacionado con aspectos genéticos del huésped (20), sin embargo estos no han sido ampliamente explorados.

La mayoría de estos estudios se han realizado en población caucásica y asiática, mientras que la población latinoamericana ha permanecido subrepresentada en cuanto a estudios que involucren factores genéticos del huésped relacionados con la severidad y la presencia de síntomas prolongados de la enfermedad (15) (16) Tomando en consideración que la población colombiana tiene una ancestría

heterogénea, se estima que presenta un perfil genético propio cuyas variaciones genéticas pueden estar implicadas en la severidad y desarrollo de secuelas de COVID-19.

Teniendo en cuenta estas consideraciones, en este estudio se realizó un análisis genético que involucró el desarrollo de un panel de secuenciación propio, y la identificación de la asociación de variables genéticas y clínicas con la severidad y la presencia de COVID-19 prolongado en una muestra población colombiana con diagnóstico confirmado de COVID-19. Asimismo, presentamos el desarrollo de un modelo multivariado predictivo de riesgo de severidad a la infección o de riesgo de presencia COVID-19 prolongado que incluyó factores genéticos y no genéticos.

2. JUSTIFICACIÓN:

2.1 Justificación general:

Según los reportes actuales de la Organización Mundial de la Salud (OMS) para noviembre del 2022, a nivel mundial persiste una alta tasa de infección siendo la región de las américas la más afectada, con un aumento en el número de casos del 19% (550.990 casos nuevos), y una la tasa de mortalidad del 21% estimado por las 3,821 nuevas muertes reportadas respecto al mes anterior (21). De manera similar, el impacto del COVID-19 sobre la población y los sistemas de salud ha sido de relevancia en Colombia. A noviembre de 2022 la Organización Panamericana de la Salud (OPS) ha reportado en Colombia un total de 6'311.259 casos y 141.820, muertes, con aumento de la incidencia del 0.02% con respecto al mes pasado de octubre, y una tasa de incidencia para Bogotá de 23.9 casos por cada 100.000 habitantes, un valor por encima de la media nacional de 12.5 (22) (23).

De manera interesante se ha descrito una heterogeneidad clínica atribuida a factores demográficos y genéticos. En este contexto, el sexo masculino, la edad de mayor de 60 años, la obesidad, la enfermedad cardiovascular, la hipertensión, la enfermedad pulmonar obstructiva crónica y el cáncer se consideran factores de riesgo y de impacto en la evolución clínica de la enfermedad (24). Desde la perspectiva genética diferentes estudios basados en GWAS y de aproximación de casos y controles han identificado diferentes polimorfismos en más de 1000 genes candidatos asociados a mayor riesgo de severidad y mortalidad por SARS-COV-2 (*ACE2*, *IFTM3*, *TMPRSS2*, *APOE*, *CCL2*, *DPP9*, *OAS1*, *TLR7*) (25) (18,26). La mayoría de los estudios genéticos han involucrado pacientes de las regiones del hemisferio norte, principalmente en Norteamérica y Europa (27), mientras que países Latinoamericanos, incluido Colombia tienen escasa o nula representación en el conocimiento de perfiles genómicos de susceptibilidad a la severidad clínica de COVID-19.

2.2 Justificación específica:

Teniendo en cuenta estas consideraciones este proyecto se justifica porque:

1. Hasta la fecha, no existen estudios que evalúen en Colombia perfiles de riesgo genómicos a gran escala, lo que representaría un aporte importante que contribuiría a cerrar la brecha de conocimiento actual en poblaciones subrepresentadas como la población Latinoamericana, de la cual sólo se ha completado 1 estudio que evalúa un número reducido de polimorfismos y 4 siguen en desarrollo (28) (27).

2. Existe la necesidad de priorización de recursos físicos y humanos de salud en el contexto de un país en vías de desarrollo, la identificación de factores de susceptibilidad podría constituir una herramienta útil para dicho objetivo. Esta herramienta ya ha sido implementada en la región europea a través de modelos predictivos de riesgo que involucran variables genéticas, mejorando la discriminación del modelo hasta en un 111% para identificar pacientes con mayor riesgo de severidad en comparación con factores demográficos y clínicos exclusivamente (29). Adicionalmente, con la transformación gradual de la pandemia a endemia, la aparición de secuelas cobrará especial importancia en la investigación, y determinar factores genéticos asociados a la aparición de estas podría ayudarnos a comprender la susceptibilidad individual a estos fenómenos.

4. Vinculación de la comunidad estudiantil en la conceptualización, desarrollo metodológico y análisis de resultados de proyectos de investigación que aborden temáticas de interés científico.

5. Va a permitir identificar grupos de pacientes de alto riesgo que se beneficiarán prioritariamente de vacunación y de manejo oportuno con antivirales

3. PREGUNTA DE INVESTIGACIÓN

3.1 Pregunta principal

¿Cuál es la asociación entre factores de riesgo genético y no genético con la severidad clínica de la infección por SARS-CoV-2 y la presencia de COVID-19 prolongado en una muestra de la población colombiana?

3.2 Preguntas Secundarias

1. ¿Cuál es la severidad del cuadro y el desenlace clínico de los pacientes analizados?
2. ¿Cuál es la asociación entre factores no genéticos con la severidad clínica de COVID-19 en una muestra de la población colombiana?

3. ¿Cuáles son las variantes potencialmente patogénicas en genes candidatos relacionados a severidad de COVID-19 en una muestra de la población colombiana?
4. ¿Cuáles son los parámetros genético-poblacionales de las variantes moleculares identificadas en la población de estudio?
5. ¿Cuál es la asociación entre el fenotipo clínico de COVID-19 y los polimorfismos de nucleótido simple (SNPs) relacionados con la respuesta a SARS-CoV 2 en una muestra de la población colombiana?
6. ¿Cuáles son las variables predictoras del riesgo de severidad por COVID-19 o de COVID-19 prolongado a partir de indicadores genéticos y no genéticos en una muestra de la población colombiana?

4. MARCO TEÓRICO

4.1 Contexto de la pandemia originada por SARS-CoV-2

El SARS-CoV-2 se identificó por primera vez en diciembre del 2019 en individuos de la población de la ciudad de Wuhan, provincia de Hubei, China (30). Previamente ya se habían identificado virus que producen manifestaciones clínicas similares a las evidenciadas en los pacientes afectados por COVID-19 como el SARS-CoV en 2003 y MERS-CoV en 2012. Estos virus son miembros de la familia *Coronaviridae* del orden *Nidovirales*, de la cual hasta el momento se conocen cuatro subgrupos principales: α , β , γ y δ . Los alfacoronavirus y betacoronavirus infectan solo a mamíferos causando enfermedades respiratorias en humanos y gastroenteritis en animales principalmente, mientras que los gammacoronavirus y deltacoronavirus infectan a las aves. En el grupo de los betacoronavirus altamente patógenos en humanos se encuentran CoV-OC43, SARS-CoV, MERS-CoV y el recientemente identificado SARS-CoV-2. Por su parte los otros 4 coronavirus detectados hasta la actualidad (HCoV-NL63, HCoV-229E, HCoV-OC43 y HKU1) inducen sólo enfermedades leves de las vías respiratorias superiores en personas con adecuada inmunocompetencia. La mayoría de estos virus tuvieron su origen evolutivo en murciélagos o roedores siendo los primeros considerados como sus reservorios naturales. Algunos otros mamíferos como los dromedarios, camélidos, civetas y pangolines, de los cuales se aisló una cepa con una similitud de secuencia del 99% al SARS-CoV-2, han contribuido como huéspedes intermediarios para finalmente llegar a causar infecciones humanas (4) (31) (32) (33).

A nivel molecular el virión de los coronavirus tiene un diámetro de entre 120-160 nm y presenta unas proyecciones dadas por la proteína S en forma de corona, figura por la cual se le atribuye su nombre y la cual media el proceso de unión y fusión de membranas durante la infección. El genoma del coronavirus codifica además para tres proteínas estructurales: a) las proteínas de membrana (M) que involucran los residuos de aminoácidos del 218 al 263, b) de envoltura (E) aminoácidos 74 al 109 y c) de la nucleocápside (N) residuos del 349 al 470. El virión del SARS-CoV-2 tiene un diámetro de 60-160 nm, su material genético está constituido por una hebra

monocatenaria de ARN en sentido positivo con una longitud de 29.891 pares de bases nucleotídicas. Según la alineación de secuencias del genoma se ha identificado que tiene una similitud del 79.5% con SARS-CoV-1. (33).

4.2 Epidemiología

En Colombia el primer caso de COVID-19 se reportó el 06 de marzo del 2020. Desde el inicio de la pandemia en nuestro país se han presentado cuatro grandes picos de contagio, el primero en agosto del 2020, el segundo en enero del 2021, el tercero en junio del 2021 y el cuarto en enero del 2022. Para la fecha del 30 de noviembre del 2022, se habían reportado 6'314.76962 casos confirmados, 6'140.844 casos recuperados, 141.895 casos fallecidos, con una tasa de letalidad de 2.25%, a la cual empezó a disminuir incluso hasta 1.6% posterior a la implementación de vacunas (23).

Bogotá ha sido la ciudad más afectada por la pandemia y de manera puntual, para esta región se han reportado 1'859.626 de casos confirmados, 1'821.695 de casos recuperados, 29.982 casos fallecidos y se ha estimado una tasa de ataque de 8.89% y una tasa de letalidad de 2.07%(23). En el transcurso de la pandemia en Colombia se han identificado aproximadamente 185 linajes virales, dentro de estas se encuentran Alfa, Delta, Gamma, Iota, Lambda, Mu y Ómicron.

En el inicio de la pandemia, las variantes más frecuentes en el territorio colombiano fueron Mu y Delta y posteriormente Ómicron. Particularmente en Bogotá las variantes de mayor impacto han sido Mu encontrada en un 39% de los casos, seguida de Delta con 25.3%, Ómicron con 10% y Gamma con un 6%, la variante menos frecuente corresponde a Iota con un 0.05% (23). La identificación de linajes compartidos, la presencia de diferentes clusters y una relación filogenética fuerte entre algunos países (Colombia, Chile, USA y algunos países europeos) ha permitido estimar que la llegada de estas variantes al país se produjo debido a casos importados (34). En orden cronológico, hubo 10 variantes que tuvieron presencia en nuestro país, de las cuales se resaltan las más prevalentes; durante el primer periodo de la pandemia (Febrero 26 – Agosto 10 del 2020) los linajes pango B.1, B.1.111, B.1.420, B.1.1.348. Posteriormente, en el segundo periodo de la pandemia (Agosto 11 del 2020 – Marzo 6 del 2021) se mantuvieron estos tres linajes de manera constante, sin embargo, empezó el reporte que evidenciaba ya la presencia de Gamma, Delta, Alfa y Mu en menor nivel. Durante el tercer periodo de la pandemia (Marzo 7 del 2021 – Julio 2021) las variantes Mu (48.4%), Gamma (21.4%), Lambda (4.5%) y Alfa (4.7%) predominaron en el panorama nacional. El periodo final de la pandemia (Agosto 2021 – Febrero 2022) se identificó predominantemente 83 subvariantes de Delta AY.x, 2 subvariantes Omicrón, Mu y una subvariante Lambda, finalmente llegando a desplazar Omicron a todas estas a finales de Febrero con una prevalencia del 87.6%. (22) (35).

Dentro de las variables epidemiológicas de interés se han identificado varios factores asociados con la infección incluyendo factores sociodemográficos, de vacunación, clínicos y genéticos del huésped. En Colombia, estudios previos han evidenciado varios factores sociodemográficos relacionados con la infección, dentro de ellos figuran convivir con más de 3 cohabitantes, el uso de transporte público, no estar afiliado a un seguro de salud, estrato socioeconómico y ocupación (policía, militar y bombero) (36). Otro factor clave implicado en la transmisibilidad y severidad de la enfermedad es la inmunización, proceso que inició en nuestro país en febrero del 2021, con la aplicación de la primera vacuna. Este último factor ha influenciado el curso de la pandemia permitiendo una disminución en la incidencia y mortalidad a causa de esta condición (34) (3)

Respecto a los factores clínicos, en Colombia se ha evidenciado un mayor requerimiento de hospitalización en UCI y mortalidad en pacientes mayores de 60 años, que el sexo masculino ha representado la mayor proporción de pacientes fallecidos con un 60.8% y que las comorbilidades de mayor impacto relacionadas con la severidad y mortalidad descritas han sido hipertensión arterial, diabetes mellitus, enfermedad respiratoria, renal o cardíaca, obesidad y cáncer (23). (24) (37) Estos hallazgos son similares a los reportados en otras poblaciones a nivel mundial.

Con relación a los factores genéticos, reportes previos han logrado encontrar variantes genéticas del huésped asociadas con la severidad de la infección. A nivel global se han realizado diversos estudios los cuales han permitido la identificación de loci de interés (18). En el contexto colombiano un estudio previo realizado por el Centro de Investigación de Genética y Genómica de la Universidad del Rosario que contó con nuestra participación, identificó una asociación estadísticamente significativa entre el polimorfismo *LZTFL1* rs11385942 con la severidad de la infección. En dicho estudio se analizaron 3 polimorfismos potencialmente asociados a la severidad de COVID-19, encontrándose que el rs11385942 estuvo asociado al desarrollo de cuadro clínico severo por COVID-19 en nuestra población (38). Es de esperar que otros estudios que amplíen el número de polimorfismos analizados o genes de interés, permitan la identificación de perfiles genómicos de susceptibilidad en nuestra población.

4.3 Fisiopatología de la infección por SARS-CoV-2

Se han descrito cuatro fases distintas en el proceso infeccioso; la fase de incubación, el periodo sintomático, la fase pulmonar y la fase hiperinflamatoria. La fase de incubación comienza con la interacción virus-celular. La proteína S viral se une al receptor ACE2 (expresado en el corazón, endotelio, hígado, riñón, testículos, intestinos, macrófagos, mastocitos, epitelio ciliado de la nasofaringe, pulmones, específicamente el epitelio bronquial, y los neumocitos tipo II, estos últimos concentran el 83% de estos receptores) (39) (40). Existe una mayor afinidad de la proteína S del SARS-CoV-2 a unirse con el ACE2 debido a su asociación con el

receptor TMPRSS2, prevalente en células epiteliales, esencial para la activación de la proteína S. Este proceso se logra gracias al dominio de unión al receptor (en inglés "Receptor Binding domain RBD) (40) el cual es clivado por la TMPRSS2, para permitir la posterior fusión de membranas, que lleva a la liberación del ARN viral y su posterior replicación (41).

La primera y segunda fase, conocidas como periodo de incubación y sintomático se centran en la infección del epitelio ciliado de la nasofaringe. Aquí inicia la respuesta de inmunidad innata, desencadenada al detectar ARN foráneo usando los receptores de reconocimiento de patrones (PRRs), estos promueven la activación de factores reguladores de interferón y factor nuclear- κ B, generando una respuesta antiviral y de citoquinas (principalmente IL-1B e IL - 18) y de interferón I y III (42) (43). El SARS-CoV-2 es capaz de inhibir la síntesis del interferón II y III, por efecto de sus proteínas no estructurales (NSP1, ORF6, ORF3B y N), controlando parcialmente la respuesta del hospedero (42).

El inóculo inicial, la tasa de replicación viral y la respuesta del hospedero determina el desenlace de la infección (44), pues una alta replicación viral a nivel del epitelio ciliar aumenta la probabilidad de la inoculación a nivel pulmonar vía orofaringe, infectando los macrófagos alveolares y neumocitos tipo II, lo que indica la progresión a la tercera fase, denominada pulmonar (45). Los macrófagos en las vías respiratorias bajas tienen mayor expresión de genes codificantes de quimioquinas proinflamatorias (*IL-8, IL1B, TNF-alfa, CCL2, CCL3, CCL5, CDCL1, CXCL3, CDCL10*), lo que explicaría la inflamación pulmonar excesiva al promover reclutamiento de monocitos y su diferenciación (41). La infección pulmonar causa daño en la integridad vascular y el intercambio gaseoso, llevando a edema pulmonar, coagulación intravascular diseminada, isquemia pulmonar, falla respiratoria hipoxémica y daño pulmonar progresivo. Adicionalmente, con un foco pulmonar establecido el virus se disemina a otros órganos incluyendo el cerebro, tracto gastrointestinal, corazón, riñón e hígado (40).

Aunque se conocen de múltiples factores que influyen el desarrollo de la fase pulmonar en COVID-19, se han identificado dos procesos causales del fenotipo de COVID-19 severo; el síndrome de activación de macrófagos alveolares y la tormenta de citoquinas. La acumulación de macrófagos activados a nivel pulmonar (síndrome de activación de macrófagos alveolares) genera un estado hiper inflamatorio constante, fenómeno explicado dado el papel de los linfocitos T CD8 (+) y los NK, los cuales previenen la acumulación de macrófagos activados: estos últimos se encuentran disminuidos y exhaustos en el COVID-19 severo, por lo que se predispone además en estos casos a coinfecciones bacterianas o micosis (41).

La invasión viral genera inflamación, activación endotelial, daño tisular y una tormenta de citoquinas mediada por el factor de crecimiento vascular endotelial (VEGF), proteína quimioatrayente de monocitos 1 (MCP-1), además de la interleuquina 8 (IL-8), lo que finalmente lleva a la reducción de E cadherina en las células epiteliales,

incrementando la permeabilidad vascular y extravasación, llevando a hipotensión, disfunción pulmonar, generando el síndrome de distrés respiratorio agudo (SDRA) el cual es insignia del fenotipo clínico severo o crítico (46).

En la fase final de hiperinflamación sistémica, los marcadores de inflamación sistémicos se elevan (IL-2, IL-6, IL-7, TNF- α , PCR, Ferritina y el Dímero D). La interacción inmunovascular genera los estados de shock, vasoplegía, insuficiencia respiratoria y el colapso cardiopulmonar (doi: 10.1016/j.healun.2020.03.012) . Es la prevalencia de esta fase, a través de la disregulación de citoquinas, activación de macrófagos y alteración de la resupuesta de las células NK, lo que finalmente lleva al cuadro crítico de infección por COVID-19 que conlleva el ingreso a unidades de cuidado intensivo (47).

4.4 Caracterización clínica de la enfermedad

Una vez un individuo es expuesto al COVID-19, el período de incubación puede durar hasta los 18 días, siendo la media de la incubación de 4 a 5 días (48) (49), no obstante, se ha visto que puede llegar a ser de 3 días dependiendo de la variante específica; La variante alfa presenta un periodo de incubación de 5 días (95% IC, 4.94 -5.06), Beta de 4.5 días (95% IC, 1.83 – 7.17), Delta de 4.41 días (95% CI, 3.76 - 5.05) y Omicrón de 3.42 días (95% IC, 3.76-5.05 días) (48). Pacientes sintomáticos presentan en orden de frecuencia tos o disnea (69.7%), fiebre subjetiva (43.1%), mialgias (36.1%), cefalea (34.4%), odinofagia (20%), diarrea (19%), y otros tales como náuseas/vómito (11.5%), anosmia/ageusia (8.3%), dolor abdominal (7.6%) y rinorrea (6.1%) (50) (51).

Es de aclarar que los signos y síntomas presentes han mostrado un cambio en su frecuencia dependiente de la variante causal; las variantes alfa y delta mostraron los OR más altos de todos para anosmia (OR 49.7, 95% IC, 44.3-55.7) y ageusia (OR 35, 95% IC, 31.9 – 40.4). Por otro lado las variantes pango BA.1 y BA.2 mostraron más frecuentemente síntomas parecidos a la gripe (“flu-like”), como fiebre (OR 18.4, 95% IC, 16.5 – 20.5). Finalmente, la variante omicrón mostró con mayor frecuencia la manifestación de síntomas de compromiso sistémico, como mialgias, artralgias, fatiga y disnea (OR 1.54, 95% IC, 1.16 – 2.06) (52)

Los casos se pueden clasificar en enfermedad asintomática (desde el 0.25% llegando hasta el 4.5% dependiendo del subgrupo poblacional analizado) (53), enfermedad leve a moderada (81%) cuyos pacientes pueden desarrollar síntomas autolimitados hasta una neumonía leve (confirmada por radiografía de tórax y saturación de oxígeno al aire ambiente >93%), sintomatología severa (14%) (presencia de disnea, saturación de oxígeno al aire ambiente <93%, o más del 50% de compromiso pulmonar en imágenes diagnósticas) y crítica (5%) (presencia de insuficiencia respiratoria, shock, sepsis o falla multiorgánica) (54) (55) Frente a los tiempos de

progresión de la enfermedad se ha observado una media de 5-8 días desde el inicio de los síntomas a la presencia de disnea, de igual manera una media de 8-12 días desde el inicio sintomático a la presencia de síndrome de distrés respiratorio agudo o admisión a UCI (50) (55).

La progresión de enfermedad a un estadio grave o crítico puede generar complicaciones, entre ellas SDRA (3.4 - 19%), arritmias (16%), lesión miocárdica (7.2% - 33%), disfunción hepática (28 - 30%), hiperglucemia (35%), rabdomiólisis (0.2%) y coagulación intravascular diseminada (0.1%). Además de coinfecciones, como las asociadas al ventilador mecánico (20%) y bacteriemia por infección de vías urinarias (3%) (55)

Se han identificado múltiples factores de riesgo clínicos para un mal pronóstico en la infección por SARS -COV-2 incluyendo presencia de disnea, frecuencia respiratoria >30 rpm, saturación de oxígeno < 93%, PaO₂/FiO₂ < 300 y/o infiltraciones pulmonares >50% del campo pulmonar. Al igual que factores de riesgo paraclínicos evidenciados en análisis univariados y multivariados, tales como los niveles de Troponina I, creatinina, ferritina sérica, lactado deshidrogenasa, interleuquina 6, dímero D, linfopenia, Escala SOFA (del inglés "Sepsis related organ failure assesment") >2, hipoalbuminemia y niveles altos de proteína C reactiva. (55)

4.5 COVID-19 prolongado

La condición de COVID-19 prolongado o post-COVID-19 es un término que desde su aparición ha recibido varias definiciones y cuyo consenso respecto a tiempo o sintomatología no ha sido claramente establecido. Esta entidad incorpora a una amplia variedad de síntomas clínicos de diversa índole e intensidad recurrentes o continuos que se presentan posterior a la infección por SARS-CoV-2. Estos pueden permanecer desde pocas semanas hasta años posterior al inicio de la enfermedad, sin embargo, en la mayoría de los casos mejora con el transcurrir del tiempo. Según la OMS el COVID-19 prolongado se define como "la enfermedad que se presenta en personas que tienen antecedentes de infección probable o confirmada por SARS-CoV-2; generalmente dentro de los tres meses desde el inicio de COVID-19, con síntomas y efectos que duran al menos dos meses". Los síntomas y efectos de la condición post COVID-19 no pueden explicarse con un diagnóstico alternativo (56). Para los Centros de Control y Prevención de Enfermedades de Estados Unidos, inglés CDC, este evento se ha denominado condición post-COVID y se define como "problemas de salud que persisten más de cuatro semanas después de haberse infectado con COVID-19" dentro de estos se incluyen:

- COVID-19 prolongado o síndrome post-COVID persistente (PPCS) en donde los pacientes tienen síntomas clínicos que pueden durar de semanas hasta meses o incluso de manera indefinida en el tiempo.

- Efectos multiorgánicos del COVID-19 haciendo referencia a manifestaciones clínicas relacionadas con los sistemas de órganos cardiovasculares, pulmonares, renales y neuropsiquiátricos
- Efectos del tratamiento/hospitalización por COVID-19, incluyendo el síndrome de cuidados post-intensivos, debilidad extrema y trastorno de estrés postraumático

Dado la integralidad del concepto para este trabajo se tomará como referencia la definición propuesta por los CDC de EE.UU. para el análisis de los aspectos relacionados con COVID-19 prolongado.

Varios estudios epidemiológicos han analizado la prevalencia de este evento encontrando una frecuencia media de 72.5% de individuos con al menos 1 síntoma post COVID (57), Carfi et al (9). En una cohorte de población italiana se describió una prevalencia del 87.4%, en Francia Carvalho-Schneider et al (58) reportó una prevalencia del 66%, en Reino Unido Arnold et al (59) reportó una prevalencia del 74%, en España Moreno-Pérez et al (60) reportó una prevalencia del 50.9% y en China Huang et al. (61) reportó una prevalencia del 76%.

Dentro de las secuelas más comúnmente identificadas se encuentran fatiga, el dolor articular o muscular, manifestaciones respiratorias como disnea y tos, cardiovasculares como palpitaciones y dolor torácico, neuropsiquiátricas como ansiedad, depresión, cefalea, alteración del sueño, ageusia y anosmia. Dentro de las secuelas menos frecuentes evidenciadas en estos estudios se presentaron algunas dermatológicas como pérdida de cabello o rash cutáneo y gastrointestinales como diarrea (62).

La fisiopatología subyacente a esta entidad no ha sido claramente definida sin embargo se reconoce su carácter multifactorial influenciado por la toxicidad viral directa, y el daño consecuente a la desregulación del sistema inmunitario y estimulación de un estado proinflamatorio en el huésped. Esto a su vez depende de varios elementos presentando diferencias interindividuales, mediadas en parte por variantes moleculares en genes del sistema inmunitario (63). Teniendo esto en consideración, se han estudiado diversos factores de riesgo implicados en el desarrollo de COVID-19 prolongado como severidad de la infección, edad 39-45 años, anormalidades pulmonares previas y sexo femenino (64). De manera similar se han descrito también algunos factores predictores de COVID-19 prolongado, como edad y número de síntomas en la primera semana de infección (65).

Dada la heterogeneidad derivada de los estudios del COVID-19 prolongado reportados a la fecha, basamos nuestra clasificación de secuelas según lo descrito y estandarizado por López-León y colaboradores (66), quienes organizaron y evaluaron las secuelas según su prevalencia y frecuencia. En este contexto y considerando los sistemas afectados, estas se pueden clasificar en: A) secuelas comunes que

incluyeron fatiga, cefalea insomnio, odinofagia, alopecia, pérdida de peso y diarrea. B) secuelas respiratorias como tos y disnea. C) secuelas cardiovasculares como dolor torácico y palpitaciones D) secuelas neurológicas como confusión, pérdida de la memoria, vértigo, anosmia, ageusia, parestesias, parálisis facial, temblores, visión borrosa. E) secuelas psiquiátricas como ánimo deprimido. F) secuelas osteomusculares tales como lumbalgia, dolor articular, mialgias, artralgias o pérdida de fuerza, y otras, entre las que se incluyeron las manifestaciones dermatológicas.

4.6. Factores clínicos

Se han identificado diferentes características que han demostrado tener significancia estadística frente al desarrollo de cuadros graves de COVID-19 entre las que se tienen variables demográficas como edad avanzada, definida como pacientes mayores de 60 años (OR: 1.93, 95% CI; 1.26 - 2.94) (67), al igual que el sexo masculino (IRR: 1.77, 95% CI; 1.74 - 1.79) (68)

Por otra parte, se han identificado condiciones preexistentes y comorbilidades relacionadas a desenlaces desfavorables en COVID-19, incrementando el riesgo de mortalidad y necesidad de intubación, tales como: hipertensión (OR 2.46, 95% CI; 1.46 - 3.83), enfermedad cardiovascular (OR 3.42, 95% CI; 1.88 - 6.22) (69), diabetes mellitus tipo II (RR 2.26, 95% CI; 1.47 - 3.49) (70), enfermedad renal crónica (OR 2.86, 95% CI; 2.47 - 3.32), enfermedad pulmonar obstructiva crónica (OR 3.34, 95% CI; 2.15 - 5.18) (71), cáncer (OR 2.73, 95% CI; 1.73 - 4.32), enfermedad cerebrovascular (2.74, 95% CI; 1.59 - 4.74) (72), antecedentes psiquiátricos (OR 1.38, 95% CI; 1.15 - 1.65) (73), demencia (OR 5.17, 95% CI; 2.31 - 11.59) (74), obesidad (OR 2.31, 95% CI; 1.3 - 4.12) (75) y tabaquismo (OR 1.91, 95% CI 1.42 - 2.59) (76)

Adicionalmente, existen otras características del individuo que han demostrado evidencia mixta y de menor calidad que justifica su asociación a mayor severidad y mortalidad por COVID-19. Entre ella está el VIH, en donde diferentes estudios han encontrado una asociación marcada dada por un hazard ratio (HR) 2.14 (95% CI; 1.70 – 2.70; $p < 0.01$) hasta un HR 2.90 (95% CI 1.96–4.30; $p < 0.0001$) (77) (78),. No obstante, una revisión sistemática que evaluaba esta patología no encontró ninguna relación (79). Adicionalmente, se han generado dudas acerca del papel de las inmunodeficiencias primarias, dado que pocos estudios tocan esta temática, además de ser más prevalente en la población pediátrica. Según datos del CDC se ha encontrado una fuerza de asociación con mortalidad dada por una razón ajustada de prevalencia (aPR) del 14.44 (95% CI; 0.99 – 29.86), no obstante, recalcan que la evidencia es insuficiente para determinar la asociación entre esta y la severidad o la necesidad de hospitalización (50). Por otra parte, el uso de corticoides como factor de riesgo para el desarrollo de fenotipos de COVID-19 severo tiene evidencia específica en pacientes con síndrome de colon irritable con aOR de 6.9 (95% CI; 2.3 – 20.5, $p < 0.01$) (80). Sin embargo, otros estudios han enfatizado la falta de asociación de terapia inmunosupresora en general con el riesgo de severidad o infección por

COVID-19, dada por una tasa de incidencia general de 0.58% (95% CI; 0.56 – 0.60) (81). Por ende, es necesario tener cautela al interpretar e incluir de manera adecuada estas variables en estudios que evalúen la severidad. Dados los resultados discrepantes, nuestra investigación incluyó estas variables para evaluar su comportamiento en la población analizada

4.7 Factores virales y moleculares

Entre las variantes de importancia clínica reconocidas desde el inicio de la pandemia, organizadas de manera cronológica se encuentran las variantes de preocupación (VOC en inglés), entre ellas: Alfa/GRY (linaje B.1.1.7), secuenciada por primera vez en septiembre del 2020 en Reino Unido. La variante Beta/GH501Y.V2 (Linaje B.1.351) la cual surgió en octubre del 2020. La variante Gamma/P1, detectada en enero del 2021. Delta (Linaje B.1.617.2), identificada en abril 2021 en India. Y por último, Ómicron (Linaje B 1.1.529) la cual fue identificada en noviembre 2021 en Sudáfrica (19,82,83)

Adicionalmente, la OMS ha identificado 8 variantes de interés, caracterizadas por presentar cambios genéticos asociados a mayor transmisibilidad, virulencia y evasión inmunológica innata y humoral (Epsilon / CAL.20C (Linajes B.1.427 y B.1.429, Eta (B.1.525) y Iota (B.1.526)). Entre las variantes en monitorización se resalta Mu (Linaje B.1.621), detectada por primera vez en enero del 2021 en Colombia, la cual presenta mutaciones importantes permitiendo una reducción de sensibilidad a los anticuerpos generados por la inmunidad natural y la vacunación. Algunas de estas mutaciones se encuentran en otras variantes de preocupación como Beta y Gamma, aumentando su transmisibilidad e infectividad, más no su severidad (84).

En Colombia, los linajes prevalentes durante la pandemia han seguido un comportamiento similar a los presentados a nivel mundial; siendo el linaje Mu (subvariante B.1.111) predominante desde mayo 2020 hasta enero del 2021, en donde pasó a ser predominante su subvariante B.1.621, al igual que Gamma (P1) y Delta (AY.25). De igual manera la variante Mu en Colombia no mostró un aumento en la mortalidad intrahospitalaria (85). Este comportamiento fue persistente hasta septiembre del 2021, en donde el linaje B.1.617.2 empezó a ser el más prevalente y responsable de más del 90% de los casos registrados para diciembre del 2021. Posteriormente y según información suministrada por Ministerio de Salud Colombiano y el Instituto Nacional de Salud de enero a mayo del 2022, Ómicron (linaje BA.1 y BA 1.1) constituyó la variante dominante en el territorio nacional.

4.8 Factores Genéticos del huésped

La severidad del COVID-19 se comporta como un rasgo multifactorial, con alta variabilidad interindividual, atribuida en parte a factores genéticos de cada individuo. Se han reportado numerosos SNPs y genes potencialmente relacionados con la severidad de la infección, incluyendo requerimiento de hospitalización, falla respiratoria, requerimiento de ventilación mecánica, falla multisistémica y muerte (14). La predisposición genética del huésped y su impacto en el fenotipo clínico de la enfermedad ha despertado gran interés en la comunidad científica de todo el mundo. Con el fin de abordar esta problemática se constituyó la iniciativa de Genética del Huésped COVID-19 (COVID-19HGI) (27), un proyecto internacional con el objetivo de identificar los determinantes genéticos del huésped implicados en la susceptibilidad y gravedad de la infección por SARS-CoV-2. Este y otros proyectos han utilizado diversas metodologías incluyendo secuenciación de siguiente generación (NGS), y estudios de asociación de genoma completo (GWAS) para identificar variantes genéticas raras y comunes, potencialmente asociadas al desenlace clínico de la entidad (81)

Los estudios de GWAS han posibilitado la caracterización de regiones genómicas, así como polimorfismos genéticos involucrados en la fisiopatología de la enfermedad. Se han descrito al menos 15 regiones genómicas relacionadas con la severidad de la enfermedad dentro de ellas destacan 3p21 (*SLC6A26*, *LZTFL1*, *CXCR6*, *CCR9*, *CCR1*, *CCR3*), 6P21.1 (*FOXP4*), MHC (*CCHCR1*, *HLA-C*), 12q24 (*OAS1*, *OAS2*, *OAS3*), 17q21.31 (*KANSL1*, *WNT3*), 17q21.33 (*TAC4*, *KAT7*), 19p13.3 (*DPP9*), 19p13.2 (*TYK2*, *ICAM1*, *ICAM3*, *ICAM4*, *ICAM5*) y 21q22 (*IFNAR2*, *IL10RB*, *TMPRSS2*)(86) El estudio más grande publicado hasta el momento que recopila los resultados de las aproximaciones genéticas desarrolladas, corresponde a un metaanálisis que incluyó 46 estudios de 19 países involucrando a un total de 49.562 casos y 2 millones de controles, este reporte tuvo en consideración individuos de diferentes ancestrias incluyendo Sur Asia (SAS) 6.489, Asia del este (EAS) 5.006, Medio este (MID) 8.876, África 8.876, América mezclado (AMR) 12.841 y Europa (EUR) 139.918 (19). En este estudio se identificaron 13 regiones genómicas independientes la mayoría de estas con variantes de riesgo localizadas en o cerca de genes del sistema inmunológico como *IFNAR2*, *CXCR6*, *LZTFL1*, *OAS1*, *OAS2*, *OAS3*, *CCR9*. Algunos de los genes candidato más estudiados a lo largo de la pandemia dado su importancia biológica en el proceso patológico son *ACE2*, *TMPRSS2*, *HLA*, *TNF- α* , *FURIN* y *ACE*, la asociación de ciertas variantes en estos genes ha sido demostrada en varios estudios Por ejemplo, la asociación de *ACE* rs4646994 el cual corresponde a una inserción deleción de un elemento alu de 287 pb con la severidad de la enfermedad, se ha documentado en 11 estudios diferentes donde el alelo deleciónado ha mostrado un incremento en el OR para fenotipo severo (OR: 1.45; 95% CI: 1.26–1.66) (87)

La asociación con severidad clínica más fuertemente establecida corresponde al locus 3p21.31 en el cual se hallan 6 genes candidatos potencialmente involucrados en el proceso fisiopatológico de la enfermedad. La variante rs11385942, localizada en el intrón 5 de *LZTFL1* se ha asociado con fenotipo severo y requerimiento de ventilación mecánica en varios estudios previos, se ha reportado un OR ajustado por sexo y edad de 2.79 (18). *LZTFL1* es un gen expresado en la vía respiratoria y está involucrado directamente con la función ciliar y ciliogénesis (88) (89,90).

Adicional a las variantes genéticas comunes identificadas a través de GWAS, las variantes genéticas raras también tienen impacto en la severidad de la enfermedad, y contribuyen a la explicación de los desenlaces clínicos. Estudios recientes han detectado este tipo de variantes en genes relacionados con mecanismos biológicos de respuesta inmune, metabolismo de los carbohidratos y reparación del ADN (*FN1*, *TLR4*, *IL17RC*, *RNASEL*, *G6PD*, *PLCG2*, *MSH6*, *STAT3*, *POLD1*). Otros estudios han identificado variantes raras de impacto clínico tales como: *PLAU* c.970 + 1G>A, *G6PD* p.Asp312His, *MYO5B* p.Ala1281Ala, *POLD1* p.Ala127*, *IL17RC* p.Arg378*, *MSH6* p.Arg1076Cys, *TLR7* p.Lys684*. y (91). *RFX2* c.899+10G>C. Esta última es de gran interés ya que ha sido asociada con la severidad de la enfermedad con un OR de 81.60 (92). Otros enfoques de análisis de variantes raras como el test de carga han favorecido la identificación de variantes y genes determinantes para la severidad de COVID-19 como en el caso del gen *TLR7* en el cual la presencia de variantes raras patogénicas se ha asociado con un incremento de 5.3 de la severidad de la enfermedad ($p=5.41 \times 10^{-7}$) (93). Dichas observaciones soportan la utilidad de análisis genómicos que identifiquen variantes genéticas raras y comunes para la identificación de factores genéticos de susceptibilidad y población de riesgo.

4.9 Estudios genéticos en Colombia

Pocos estudios de perfiles genómicos de susceptibilidad relacionados con la severidad clínica de COVID -19 han sido desarrollados en el país. Recientemente, nuestro grupo de investigación publicó un análisis de asociación caso-control entre el desenlace de COVID-19 y tres SNPs (rs11385942, rs4646994 y rs2285666) ubicados en los genes *LZTFL1*, *ACE* y *ACE2*. Este estudio identificó que la variante *LZTFL1* rs11385942 ($p < 0.01$; OR = 5.73; 95% CI = 1.2–26.5), está asociada a la severidad de la enfermedad. Sin embargo, y en discordancia con otros estudios, nuestros análisis no encontró asociación con las variantes genéticas *ACE* rs4646994 y *ACE2* rs2285666 (38).

4.10 Herramientas de priorización en Salud pública y Modelos de regresión aplicados a COVID19

Diferentes instrumentos para evaluar la priorización de atención de pacientes con COVID-19 se han desarrollado a través del transcurso de la pandemia enfocados principalmente en pacientes hospitalizados y en observaciones clínicas. Una de las primeras herramientas extrapoladas para priorización de pacientes COVID-19 fue la escala nacional de signos de alarma en inglés National Early Warning Score 2 (*NEWS-2*), que determina el grado de enfermedad de un paciente para realizar intervenciones en cuidado crítico. Esta escala presenta una sensibilidad del 98% (95% CI; 0.96 - 1.00) y especificidad del 28% (95% CI; 0.21 - 0.35), con un valor predictivo negativo (VPN) de 96 (95% CI; 0.90 - 1.00) y un área bajo la curva ROC (AUC ROC) de 0.71 (95% CI; 0.65 - 0.77, $p < 0.001$).

Por otra parte, han aparecido escalas de evaluación primaria, cuyo objetivo es usar datos paraclínicos no invasivos, con ayuda de la anamnesis, capaces de evaluar severidad por COVID-19 al momento de la admisión hospitalaria. Tal como es el caso de Covichem, que al utilizar un modelo de regresión logística permite la obtención de un score con una sensibilidad del 80% y especificidad del 85%, utilizando como base variables clínicas y paraclínicas sencillas (obesidad, enfermedad cardiovascular, natremia, albúmina, ferritina, LDH y CK) (94).

Otra herramienta, el COVID-19 severity index (*qCSI*) genera un modelo con AUC ROC de 0.76 (95% CI; 0.65 - 0.86%), sensibilidad de 73% (95% CI; 0.56 - 0.88) y especificidad de 81% (95% CI; 0.72 - 0.89) (95). Estas estrategias han surgido ante la necesidad de crear modelos para evaluar la severidad de la enfermedad y ser utilizadas en las diferentes etapas de la evolución clínica, ya sea triage, periodo de hospitalización ó pronóstico en Unidad de Cuidados Intesivos (UCI).

Aún así la aplicación en salud pública de estas herramientas es limitada, frente a este contexto, la utilización de biomarcadores adicionales a factores clínicos y paraclínicos rutinarios ha tomado cada vez más importancia, es por esto que la inclusión de marcadores genéticos, a través de la identificación de polimorfismos de riesgo ha sido relevante, novedoso y puede significar una solución a largo plazo como instrumento de priorización de atención en salud.

Actualmente modelos que combinan estos tres componentes han tenido buenos resultados. Al construir modelos predictivos usando la regresión logística, se han logrado AUC ROC de 0.732, (95% CI; 0.78 - 0.756 (96) y de 0.635; 95% CI; 0.607 - 0.662), los cuales obtuvieron un mayor poder discriminatorio hasta de 111% en comparación con modelos exclusivamente clínicos, al incluir dentro de ellos polimorfismos genéticos asociados a la severidad de COVID-19 (29). Países como Holanda y Australia han integrado holísticamente los componentes clínicos y genéticos. En la actualidad en Colombia, no existen modelos predictivos extendidos que incluyan varios predictores genéticos.

5. HIPÓTESIS

5.1 Hipótesis nula

El desenlace clínico y la persistencia de manifestaciones clínicas en las personas afectadas por COVID-19 no está asociado con factores clínicos y variantes genéticas y no determina perfiles de susceptibilidad individuales en la severidad de la enfermedad.

5.2 Hipótesis alterna

El desenlace clínico y la persistencia de manifestaciones clínicas en las personas afectadas por COVID-19 está asociado con factores clínicos y variantes genéticas, lo que determina perfiles de susceptibilidad individuales en la severidad de la enfermedad.

6. OBJETIVOS

6.1. Objetivo general

Determinar la asociación entre factores de riesgo genéticos y no genéticos con la severidad clínica de COVID-19 y la presencia de COVID-19 prolongado en una muestra de la población colombiana.

6.2. Objetivos específicos

1. Caracterizar el papel de variables demográficas y clínicas con la severidad del cuadro y el desenlace clínico de los pacientes analizados.
2. Determinar los parámetros genético-poblacionales de las variantes moleculares identificadas en la población de estudio
3. Identificar variantes moleculares potencialmente patogénicas en genes candidatos asociados a severidad de COVID-19 en una muestra de la población colombiana
4. Estimar la asociación entre factores de riesgo no genético con la severidad clínica de COVID-19 o con la presencia de COVID-19 prolongado en una muestra de la población colombiana
5. Establecer la asociación entre el fenotipo clínico o la presencia de COVID-19 prolongado y polimorfismos en genes candidatos relacionados con la respuesta a SARS-CoV-2 en una muestra de la población colombiana
6. Proponer un modelo predictivo para evaluar el riesgo de severidad por COVID-19 o la presencia de COVID-19 prolongado a partir de indicadores genéticos y no genéticos en la población colombiana.

7. METODOLOGÍA

7.1 Enfoque metodológico

El enfoque metodológico del presente estudio corresponde a un estudio cuantitativo de casos y controles retrospectivo observacional analítico. Los casos fueron definidos como pacientes entre los 18 – 60 años con resultado RT – PCR, antígeno (Ag) o anticuerpo (Ac) nucleocápside positivos para SARS-CoV-2 y que hayan cumplido los criterios de inclusión. Y que sean considerados casos moderados, severos o críticos según el consenso de la asociación Colombiana de infectología (ACIN) y el Instituto nacional de salud (INS) para el 2020 (51). El muestreo usó una aproximación no probabilística, e incluyó un total de 112 pacientes, en relación de casos y controles de 1:1. Los pacientes fueron reclutados entre 11/12/2020 y 07/07/2021. Todos los individuos incluidos en el estudio no se encontraban vacunados o solo contaban con una dosis administrada durante los 7 días previos al inicio de los síntomas. Teniendo en cuenta el impacto clínico de la edad en la evolución y desenlace del COVID-19 en este estudio se parearon casos y controles por edad. Se recolectó información sociodemográfica, clínica y genética de los pacientes, y se analizó su asociación con la severidad clínica de COVID-19 y la presencia COVID-19 prolongado.

7.2 Tipo de estudio

Estudio observacional, analítico, retrospectivo en una población con evidencia serológica o por PCR de infección por SARS-CoV-2.

7.3 Población

7.4 Descripción de población de estudio

La población de análisis del estudio estuvo conformada por individuos de nacionalidad colombiana entre 18 y 60 años, con prueba RT-PCR, antígenos y/o anticuerpos de nucleocápside positivos para SARS-CoV-2 que decidieron participar voluntariamente en el estudio sin historial de vacunación o sólo una dosis administrada durante los 7 días previos al inicio de síntomas. Los pacientes se dividieron en dos grupos definidos como casos y controles. Los casos correspondieron a pacientes con cuadro clínico severo o crítico. Este grupo estuvo conformado por individuos que secundario al COVID-19, requirieron hospitalización y/o oxigenoterapia, y pacientes que requirieron de manejo en UCI, respectivamente. La definición de la severidad se realizó según los lineamientos para el manejo clínico de pacientes con infección por SARS-CoV-2 del Ministerio de Salud (PSSS03 v02). Estos pacientes se identificaron y reclutaron exclusivamente en el Hospital Universitario Mayor Méderi en el servicio de Urgencias y hospitalización del hospital.

Los controles correspondieron a pacientes asintomáticos y con síntomas leves. Este grupo estuvo conformado por pacientes que no manifestaron sintomatología relacionada con COVID-19 o que tuvieron indicación de manejo domiciliario, sin requerimientos de hospitalización a causa de la enfermedad. Los casos asintomáticos

o leves se identificaron y reclutaron en el laboratorio de Genética Molecular de Colombia.

7.5 Definición de la unidad de análisis:

Se define como unidad de análisis el paciente que cumple con criterios de selección definidos para el estudio, para los que se identificó y analizó:

1. Severidad del cuadro clínico de COVID-19, sintomatología y secuelas clínicas.
2. Factores no genéticos incluyendo sexo, edad, presencia de comorbilidades, entre otros.
3. Variantes en genes y regiones candidatas.

7.6 Criterios de inclusión y exclusión

7.6.1 Criterios de inclusión generales:

- Paciente asintomático o caso leve, severo o crítico, clasificado según los lineamientos para el manejo clínico de pacientes con infección por nuevo coronavirus COVID-19 del Ministerio de Salud (PSSS03 v02) con diagnóstico confirmado de infección por SARS-CoV-2 mediante prueba de antígenos, RT-PCR o anticuerpos nucleocápside.
- Entre 18 y 60 años.
- Nacionalidad colombiana.
- Participación voluntaria en el estudio

7.6.2 Criterios de inclusión grupo casos

Paciente sintomático clasificado como severo o crítico según la Guía de Ministerio de Salud de 2020 con COVID-19 confirmado por prueba de RT-PCR, antígenos y/o anticuerpos (51). Los pacientes de este grupo se identificaron e incluyeron en el Hospital Universitario Mayor Méderi. Los pacientes severos son aquellos que presentaron: disnea, frecuencia respiratoria ≥ 30 /min, saturación de oxígeno en sangre $\leq 93\%$, Índice de Kirby o $\text{PaO}_2/\text{FiO}_2 < 300$ y/o infiltrados pulmonares $> 50\%$ en 24 a 48 horas. Por su parte se consideraron pacientes críticos aquellos con: insuficiencia respiratoria, choque séptico y/o disfunción o falla de múltiples órganos. Una vez la IPS realizó el ingreso hospitalario, profesionales adscritos al proyecto de investigación invitaron a la participación en el estudio y los voluntarios firmaron el consentimiento informado. Posterior a la inclusión voluntaria se hizo un seguimiento telefónico para verificar la condición clínica del paciente por un periodo de 30 días luego de la toma de muestra y se confirmó en las historias clínicas el cumplimiento de los criterios de severidad.

7.6.3 Criterios de inclusión grupo controles:

Pacientes asintomáticos (sin tos, fiebre, anosmia, odinofagia u otros síntomas) o con sintomatología clasificada como leve, que no requirieron hospitalización, con diagnóstico de COVID-19 confirmado por prueba RT-PCR, antígenos y/o anticuerpos de nucleocápside, y sin cambios del desenlace clínico 14 días después de la toma de muestra. Las personas asintomáticas o leves incluidas fueron exclusivamente aquellas que solicitaron la realización de dicha prueba en el laboratorio de Genética Molecular de Colombia (GMC). Se consideraron casos leves aquellos que no presentan neumonía o con neumonía leve sin requerimiento de hospitalización. Una vez la institución prestadora de servicios de salud (IPS) GMC atendió el caso, los profesionales de salud de esta institución adscritos al proyecto invitaron a las personas con resultado positivo a participar en el estudio. Posterior a la inclusión voluntaria se realizó un seguimiento telefónico para verificar que estas personas inicialmente asintomáticas o con síntomas leves mantuvieron su condición clínica estable por un periodo de 21 días luego de la toma de muestra.

7.6.4 Criterios de exclusión

- Pacientes que no acepten participar en el estudio de manera verbal o escrita durante cualquier fase de realización del estudio.
- Pacientes con historial de vacunación o con una dosis administrada más de 8 días previos al inicio de síntomas.

7.7 Tamaño de la muestra

7.7.1 Método de muestreo

Para el estudio se eligió un muestreo no probabilístico puesto que no se conoce *a priori* la probabilidad que un paciente que acuda a Méderi por sospecha de COVID-19 realmente forme parte de la muestra y la selección de los pacientes depende de los criterios de inclusión y de exclusión. Por esta razón, la muestra se seleccionó mediante un proceso no aleatorio, ya que se incluyeron los pacientes que cumplieron los criterios de inclusión y de exclusión y que aceptaron voluntariamente participaren el estudio.

7.7.2 Cálculo del tamaño de la muestra

Se definió un tamaño muestral de acuerdo con la frecuencia alélica mínima de la variante genética rs11385942 en el gen *LZTFL1*, la cual se encuentra en el locus 3p21.31, identificada como la región genómica asociada con severidad clínica con mayor evidencia de literatura. De esta manera se calculó el tamaño de la muestra usando la fórmula $n = \frac{NZ^2p(1-p)}{e^2(N-1)+Z^2p(1-p)}$ implementada en la herramienta web de OpenEpi, se usó una proporción (p) de 5% de acuerdo a la frecuencia alélica mínima para esta variante obtenida de un estudio previo en población Colombiana (38), un intervalo de confianza del 95% ($\alpha=0.05$, $Z=1.96$), un error (e) del 5% y un tamaño poblacional finito $N = 8.000.000$ para la ciudad de Bogotá, el valor estimado es de 73 individuos. El tamaño muestral preliminar definido a conveniencia fue determinado

tras la aplicación de los criterios de inclusión y exclusión, obteniéndose un número de muestras de 119. Respecto a la potencia estadística al tamaño muestral estimado es del 98% esta fue calculada usando la herramienta web https://www.statskingdom.com/32test_power_t_z.html con el OR obtenido del mismo estudio previo en población Colombiana (38).

8. Definición de las variables:

La siguiente tabla ilustra la definición, escala de medición, tipo de variable, relación y objetivo diana de cada variable. La definición de cada una de las variables fue extraída y pulida basada en dos sistemas de búsqueda: los DECS (<https://decs.bvsalud.org/es/>) y el diccionario de la real academia española (RAE) (<https://dle.rae.es/diccionario>)

Tabla 0: Definición de las variables

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Edad	Número de años cumplidos a la fecha de la entrevista	Años	Razón	OBJ 5-6
Sexo	Condición biológica que diferencia a los hombres y mujeres	Femenino Masculino	Nominal	OBJ 5-6
Días de seguimiento	Números de días entre la inclusión en el estudio y de seguimiento clínico	Días	Razón	OBJ 1
Fecha de diagnóstico	Día de obtención de prueba con resultado positivo para SARS-CoV2	Días	Razón	OBJ 1
Fecha de inicio de síntomas	Día de inicio de la sintomatología asociada a la infección por COVID19	Días	Razón	OBJ 1
Tipo de prueba diagnóstica	Clase de herramienta inmunológica o molecular utilizada para confirmar la sospecha diagnóstica por COVID19.	RT-PCR Antígenos Anticuerpos	Nominal	OBJ 1
Grupo Sanguíneo	Principal sistema de tipos sanguíneos humanos que depende de la presencia o ausencia de dos antígenos A y B. El tipo O ocurre cuando no está presente ni el A ni el B y AB cuando ambos estén presentes.	Letras y signos	Nominal	OBJ 4 y 6

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Depresión	Un trastorno afectivo que se manifiesta tanto por un estado anímico disfórico como por la pérdida de interés o placer en las actividades usuales. El trastorno del ánimo es prominente y relativamente persistente.	Si: el paciente autopercebe que presentó síntomas depresivos No: el paciente no autopercebe que presentó síntomas depresivos	Nominal	OBJ 4 y 6
Enfermedad coronaria	Denominación general para una variedad de afecciones que afectan la estructura y función del corazón reportada en Historia clínica para grupo B y reportada por paciente para Grupo A.	SI: Sufre de la patología NO: No sufre de la patología	Nominal	OBJ 4 y 6
Antecedente de Depresión	Presencia en historial clínico el trastorno anímico caracterizado por ser un trastorno afectivo que se manifiesta tanto por un estado anímico disfórico como por la pérdida de interés o placer en las actividades usuales. El trastorno del ánimo es prominente y relativamente persistente.	Si: paciente presenta diagnóstico previo de depresión No: paciente no presenta diagnóstico previo de depresión	Nominal	OBJ 4 y 6
Arritmias	Trastorno de la frecuencia o del ritmo cardíaco. El corazón puede latir demasiado rápido (taquicardia), demasiado lento (bradicardia) o de manera irregular. Reportada en Historia clínica para grupo B y reportada por paciente para Grupo A	SI: Sufre de la patología NO: No sufre de la patología	Nominal	OBJ 4 y 6

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Hipertensión arterial	<p>Presión arterial sistólica mayor a 130mmHg y/o presión arterial diastólica mayor a 80 mm Hg.</p> <p>Reportada en Historia clínica para grupo B y reportada por paciente para Grupo A</p>	<p>SI: Sufre de la patología</p> <p>NO: No sufre de la patología</p>	Nominal	OBJ 4 y 6
EPOC	<p>Enfermedad pulmonar progresiva caracterizada por una reducción persistente del flujo de aire. Reportada en Historia clínica para grupo B y reportada por paciente para Grupo A</p>	<p>SI: Sufre de la patología</p> <p>NO: No sufre de la patología</p>	Nominal	OBJ 4 y 6
Asma	<p>Enfermedad del sistema respiratorio dada por inflamación crónica de la vía aérea, cuyas manifestaciones clínicas son heterogéneas y variables en el tiempo (sibilancias, dificultad respiratoria, opresión torácica y tos). Reportada en Historia clínica para grupo B y reportada por paciente para Grupo A</p>	<p>SI: Sufre de la patología</p> <p>NO: No sufre de la patología</p>	Nominal	OBJ 4 y 6
Diabetes mellitus	<p>Trastorno metabólico caracterizado por concentraciones altas de glucosa en sangre de forma crónica. causado por un defecto en la producción de insulina, resistencia a la acción de la insulina, aumento en la producción de glucosa o una combinación. Reportada en Historia clínica para grupo B y reportada por paciente para Grupo A</p>	<p>SI: Sufre de la patología</p> <p>NO: No sufre de la patología</p>	Nominal	OBJ 4 y 6

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Enfermedad renal	Pérdida progresiva persistente de la función renal. reportada en Historia clínica para grupo B y reportada por paciente para Grupo A	SI: Sufre de la patología NO: No sufre de la patología	Nominal	OBJ 4 y 6
Cáncer	Enfermedades en las que se observa un proceso descontrolado en la división de las células del cuerpo. reportada en Historia clínica para grupo B y reportada por paciente para Grupo A	SI: Sufre de la patología NO: No sufre de la patología	Nominal	OBJ 4 y 6
Obesidad	Acumulación anormal o excesiva de grasa que puede ser perjudicial para la salud. reportada en Historia clínica para grupo B y reportada por paciente para Grupo A	SI: Sufre de la patología NO: No sufre de la patología	Nominal	OBJ 4 y 6
Trasplantes	Sustitución de un órgano o tejido enfermo por otro que funcione adecuadamente. Reportada en Historia clínica para grupo B y reportada por paciente para Grupo A	SI: Sufre de la patología NO: No sufre de la patología	Nominal	OBJ 4 y 6
Uso prolongado de corticoides	Uso de corticoides mayor a 14 días. Reportada en Historia clínica para grupo B y reportada por paciente para Grupo A	SI: Ha usado de manera prolongada corticoides NO: No ha usado de manera prolongada corticoides	Nominal	OBJ 4 y 6
VIH/ Inmunodeficiencias	Presencia de inmunodeficiencias adquiridas o hereditarias diferentes a uso prolongado de corticoides. Definida como: Estado del organismo consecuente a una deficiencia funcional del sistema	Si: Paciente con VIH/ inmunodeficiencias diagnosticado	Nominal	OBJ 4 y 6

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Enfermedades Autoinmunes	Presencia de cualquier enfermedad autoinmune como antecedente. Definida como: Trastornos que se caracterizan por la producción de anticuerpos que reaccionan con los tejidos del hospedero o con las células efectoras inmunes que son autorreactivas a péptidos endógenos.	Si: paciente con antecedente de enfermedad autoinmune No: paciente sin antecedente de enfermedad autoinmune	Nominal	OBJ 4 y 6
Convulsiones	Trastorno que se caracteriza por episodios recurrentes de paroxismos de disfunción cerebral debido a una descarga neuronal súbita, desordenada y excesiva	Si: paciente presenta convulsiones como antecedente No: paciente no presenta convulsiones como antecedente		OBJ 4 y 6
Antecedente de tabaquismo	Historial de adicción al consumo de tabaco por la nicotina en algún momento de la vida	Si: paciente con antecedente de tabaquismo No: paciente sin antecedente de tabaquismo	Nominal	OBJ 4 y 6
Tabaquismo activo	Adicción al consumo de tabaco por la nicotina en el último año.	SI: Ha sido consumidor de tabaco NO: No ha sido consumidor de tabaco	Nominal	OBJ 4 y 6
Numero de cigarrillos/día	Cantidad de cigarrillos consumidos por día	Número de cigarrillos/día	Razón	OBJ 4 y 6

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Años de tabaquismo	Número de años en los cuales el paciente consumió activamente tabaco	Años	Razón	OBJ 4 y 6
IPA	Razón entre el número de cigarrillos fumados en un día y el número 20	Numérico	Razón	OBJ 4 y 6
Antecedente de adicción a drogas	Dependencia clínica del paciente a una sustancia de abuso (Drogas recreativas o sustancias químicas)	Si: paciente presenta adicción a drogas No: paciente no presenta adicción a drogas	Nominal	OBJ 4 y 6
Antecedentes farmacológicos	Medicamentos consumidos durante el último año	Tipo de medicamento o consumidos	Nominal/N umérica	OBJ 4 y 6
Tiempo de tabaquismo	Tiempo en el que ha consumido tabaco.	Años	Razón	OBJ 4 y 6
Comorbilidades ninguna	Nula presencia de cualquier comorbilidad previamente estipulada en el protocolo	Si: paciente sin ninguna comorbilidad No: paciente con mínimo una comorbilidad	Nominal	OBJ 4 y 6
Comorbilidades 2	Presencia de 2 o más comorbilidades previamente estipuladas en el protocolo	Si: paciente con 2 o más comorbilidades estipuladas NO: paciente con menos de 2 o ninguna comorbilidad estipulada	Nominal	OBJ 4 y 6

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Disnea	Respiración difícil o laboriosa.	Si: paciente presentó disnea No: paciente no presentó disnea	Nominal	OBJ 1, 4 y 6
Tos	Expulsión súbita y audible del aire de los pulmones a través de una glotis parcialmente cerrada, precedida por la inhalación.	Si: el paciente presentó tos. No: el paciente no presentó tos	Nominal	OBJ 1, 4 y 6
Hemoptisis	Expectoración o esputo de sangre, procedentes de cualquier parte del tracto respiratorio, generalmente de hemorragias en la parénquima pulmonar	Si: el paciente presentó hemoptisis No: el paciente no presentó hemoptisis	Nominal	OBJ 1, 4 y 6
Anosmia	Pérdida o deficiencia de la capacidad olfatoria	Si: el paciente presentó anosmia No: el paciente no presentó anosmia	Nominal	OBJ 1, 4 y 6
Ageusia	Pérdida completa o severa del sentido subjetivo del gusto	Si: el paciente presentó ageusia No: el paciente no presentó ageusia	Nominal	OBJ 1, 4 y 6
Fatiga	Signos o síntomas clínicos manifestados como debilidad o falta o pérdida de fuerza y energía	Si: el paciente presentó fatiga No: el paciente no presentó fatiga	Nominal	OBJ 1, 4 y 6

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Dolor Musculoesquelético	Dolor no inflamatoria en las articulaciones y músculos	Si: el paciente presentó dolor musculoesquelético. No: el paciente no presentó dolor musculoesquelético	Nominal	OBJ 1, 4 y 6
Fiebre	Aumento de la temperatura corporal por encima de los 38°C	Si: el paciente presentó fiebre No: el paciente no presentó fiebre	Nominal	OBJ 1, 4 y 6
Diarrea	Aumento de la liquidez o disminución de la consistencia de las HECES, con deposiciones seguidas	Si: Paciente quien presentó diarrea No: Paciente que no presentó diarrea	Nominal	OBJ 1, 4 y 6
Náuseas/Emesis	Sensación desagradable en el estómago acompañada generalmente de la necesidad de vomitar / Expulsión forzada del contenido del estómago a través de la boca	Si: paciente presentó náuseas o emesis No: paciente no presentó ni náuseas o emesis	Nominal	OBJ 1, 4 y 6

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Desmayos/Desorientación/Somnolencia	Pérdida transitoria de la conciencia y del tono postural que se origina por disminución del flujo sanguíneo cerebral/Pérdida de la noción del tiempo y del espacio, propia de las alteraciones anatómicas o de la función del sistema nervioso central./Trastornos caracterizados por hipersomnolencia durante las horas normales de vigilia que pueden afectar la función cognitiva	Si: paciente presentó Desmayos, Desorientación o somnolencia No: paciente no presentó ni, desmayos, desorientación o somnolencia	Nominal	OBJ 1, 4 y 6
Congestión nasal	Sensación de taponamiento a nivel de la nariz o los senos paranasales. Acompañada o no de rinorrea.	Si: paciente presentó congestión nasal No: paciente no presentó congestión nasal	Nominal	OBJ 1, 4 y 6
Cefalea	Síntoma de dolor en la región craneal	Si: paciente presentó cefalea No: paciente no presentó Cefalea	Nominal	OBJ 1, 4 y 6
Tiempo de recuperación	Tiempo en el cual el paciente se recuperó de la sintomatología	Días	Razón	OBJ 1
Tiempo de persistencia de síntomas	Tiempo en el cual el paciente persistía con síntomas	Días	Razón	OBJ 1 y 6
Síntoma prolongado Vértigo	Presentación como Síntoma prolongado de la Ilusión de movimiento, bien del mundo externo que se mueve alrededor del individuo o del individuo que se mueve en el espacio.	Si: paciente presenta secuela de vértigo No: paciente no presenta secuela de vértigo	Nominal	OBJ 1

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Síntoma prolongado Ageusia/Anosmia	Presentación como Síntoma prolongado de la Pérdida o deficiencia de la capacidad olfatoria / Pérdida completa o severa del sentido subjetivo del gusto	Si: paciente presenta secuela de ageusia / anosmia No: paciente no presenta secuela de ageusia / anosmia	Nominal	OBJ 1
Síntoma prolongado Confusión	Presentación como Síntoma prolongado de estado mental caracterizado por confusión, trastornos emocionales, falta de pensamiento claro y desorientación perceptiva	Si: paciente presenta como secuela confusión No: paciente no presenta como secuela confusión	Nominal	OBJ 1
Otros Síntoma prolongado	Presentación como Síntoma prolongado la presencia de erupciones cutáneas o fenómenos tromboembólicos o autoinmunidad	Si: paciente presenta cualquier otra secuela No: paciente no presenta ninguna de las otras secuelas	Nominal	OBJ 1
Tratamiento	Tratamiento de una enfermedad o afección por varios medios diferentes de forma simultánea o secuencial	Tipo de medicamento o grupo de acciones guiadas por protocolos farmacológicos enfocados a tratar COVID19	Nominal	OBJ 1 y 6

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Vacuna COVID19	Sustancia realizada a través de diferentes plataformas (mRNA, virus vivo, virus atenuado) que se introduce en el organismo para prevenir y tratar determinadas enfermedades infecciosas; estimula la formación de anticuerpos con lo que se consigue una inmunización contra estas enfermedades, en este caso COVID-19	Si: Paciente si fue vacunado contra COVID19 No: Paciente no fue vacunado contra COVID19	Nominal	OBJ 1
Marca Vacuna COVID19	Laboratorio al cual se le atribuye la vacuna COVID-19 con la cual se inmunizó al paciente	SINOVAC AstraZeneca Pfizer Moderna Janssen	Nominal	OBJ 1
Signos clínicos de dificultad respiratoria	Evidencia clínica que demuestra que el paciente cursa con dificultad respiratoria tales como requerimiento de oxígeno suplementario, uso de músculos accesorios y disnea	Si: paciente presenta signos clínicos de dificultad respiratoria NO: paciente no presenta signos clínicos de dificultad respiratoria	Nominal	OBJ 1 y 6
Saturación de oxígeno por debajo de 93%	Evidencia basada en pulsioximetría que demuestra saturaciones concordantes con hipoxemia, que requieren el uso de oxígeno suplementario	Si: paciente presenta saturación de oxígeno < 93% No: paciente no presenta saturación de oxígeno < 93%	Nominal	OBJ 1 y 6

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
PAFI < 300	Razón utilizada para evaluar la gravedad de los trastornos de oxigenación y la insuficiencia respiratoria	Si: paciente que presenta insuficiencia respiratoria leve, moderada o grave. No: paciente que no presenta insuficiencia respiratoria	Nominal	OBJ 1 y 6
Infiltrados pulmonares	Presencia de signos radiológicos concordantes con infiltrados a nivel pulmonar que ocupan un porcentaje > 50% de los campos pulmonares	Si: paciente que presenta infiltrados pulmonares NO: paciente que no presenta infiltrados pulmonares	Nominal	OBJ 1 y 6
Insuficiencia respiratoria aguda	Es la incapacidad del sistema respiratorio de cumplir su función básica, que es el intercambio gaseoso de oxígeno y dióxido de carbono. Evidenciada clínicamente no sólo por parámetros bioquímicos, sino también por la necesidad de intubación del paciente	Si: paciente con insuficiencia respiratoria aguda NO: paciente sin insuficiencia respiratoria aguda	Nominal	OBJ 1 y 6
Choque séptico	Subcategoría de la sepsis en la que las alteraciones circulatorias y del metabolismo celular son lo suficientemente profundas como para aumentar considerablemente la mortalidad.	Si: paciente con presencia de shock séptico NO: paciente sin presencia de shock séptico	Nominal	OBJ 1 y 6
Variantes genéticas en genes asociados a severidad de COVID –19 (74 genes)	Constitución genética del individuo	Homocigoto alelo <i>wild type</i> Heterocigoto Homocigoto alelo polimórfico	Nominal	OBJ 3

Variable	Definición	Escala de medición	Tipo de variable	Objetivo por el cual se utiliza
Variantes genéticas asociados a severidad de COVID-19 (81 SNP)	SNP Polimorfismos de nucleótido simple	Homocigoto alelo <i>wild type</i> Heterocigoto Homocigoto alelo polimórfico	Nominal	OBJ 2,5 y 6
Severidad de COVID 19	Manifestación clínica de la infección por SARS Cov2	Asintomático Leve Severo Crítico	Nominal	OBJ 1, 4, 5 y 6
Presencia COVID-19 prolongado	Presencia de COVID-19 prolongado	Ausencia Presencia Tipo de secuela	Binario Categoríca	OBJ 1, 4, 5 y 6

7.9 Técnicas de recolección de la información

7.9.1 Fuente de información

Los pacientes que acudieron o hayan acudido a Mederi o a Genética Molecular de Colombia para toma de prueba RT-PCR, antígenos y/o anticuerpos para SARS-CoV-2 o para atención hospitalaria por enfermedad por COVID-19 y que cumplieron los criterios de inclusión se invitaron a participar en el estudio. Tras la obtención del consentimiento informado (Anexo1), se recolectaron datos del formulario clínico y datos sociodemográficos. Se explicó en detalle el objetivo de la investigación y se invitó al participante a leer el contenido del documento del consentimiento informado (el cual fue previamente aprobado por el comité de ética de la Universidad del Rosario DVO0051543-CV1334). Se reforzó la información de los potenciales beneficios del estudio y se estableció que el proyecto no contemplaba ningún beneficio económico por la participación. El equipo investigador explicó las garantías que se ofrecieron al participante en términos del anonimato en el manejo de los datos y la custodia que sobre ellos se ejercería. No se requirió la firma de un testigo, dadas las condiciones sanitarias y el potencial riesgo de contagio por COVID-19 para acompañantes de los pacientes.

Se diseñó conjuntamente entre los investigadores una encuesta telefónica que se aplicó a los usuarios de manera uniforme, con preguntas consecutivas tanto abiertas como cerradas (Anexo 2). Se incluyeron en el cuestionario las variables de identificación del individuo y las sociodemográficas. Los datos clínicos del paciente involucraron: antecedentes personales, sintomatología inicial, su cronicidad, la presencia de secuelas con sus características cronológicas y el tratamiento suministrado. Cualquier información pertinente adicional pudo ser anotada por el investigador en una sección específica titulada “comentarios” para ser considerada durante la interpretación y digitalización de las encuestas.

Para la estandarización de la encuesta se realizaron pruebas piloto para evaluar el entendimiento, la sincronía y la pertinencia de cada una de las preguntas a realizar. El sesgo de memoria e información derivada de la aplicación de la encuesta en los pacientes incluidos fue controlada mediante comparación y corrección con la historia clínica. Adicionalmente, se discutió con cada uno de los pacientes nuevamente su asentimiento para la toma de toda información sensible a través de su asentimiento verbal. Las encuestas fueron realizadas de manera telefónica, garantizando la comodidad del paciente y por ende la calidad del dato.

Las encuestas fueron digitalizadas en una base de datos creada en el programa Microsoft Excel. Se garantizó la confidencialidad de la información y el acceso restringido contemplado en las consideraciones éticas evaluadas por el comité de ética de la Universidad del Rosario.

La información de los pacientes se mantuvo en las instalaciones del Centro de Genética y Genómica de la Universidad del Rosario (CIGGUR) en archivadores debidamente cerrados y con acceso exclusivo del Investigador principal y del director del Centro de Investigación. Los consentimientos informados fueron firmados por duplicado, uno de ellos reposa en el Centro de investigación en genética y genómica de la Universidad del Rosario, bajo la custodia del grupo de investigación y la otra copia fue entregada al paciente. El equipo investigador escaneó el consentimiento y lo almacenó en carpetas aseguradas con acceso exclusivo del investigador principal. Los participantes se identificaron como: COV01 y la numeración se realizó de forma consecutiva. El dato de la información genómica de los pacientes se mantuvo anónima y se mantiene en un equipo de cómputo con acceso exclusivo del Investigador principal y del director del Centro de Investigación. Los datos se usaron para el objeto de la esta investigación.

Las muestras de ADN se extrajeron y cuantificaron en las instalaciones del CIGGUR. Estas muestras se guardaron con el código asignado de cada paciente en congeladores de -30°C destinados para tal fin. La fecha de extracción de ADN, la localización de las muestras y su cuantificación se registró en un formato de laboratorio dispuesto para tal fin al cual solo tendrán acceso los coinvestigadores asignados a este rol. Los resultados de las secuenciaciones de los genes y SNP de interés se consignaron en las bases de datos pertinentes y la información sólo se conoce por los investigadores principales. Todas las muestras fueron procesadas e identificadas con el código asignado al paciente. No se usaron nombres para mantener así la confidencialidad.

7.10 CONTROL DE SESGOS:

Sesgo de selección: derivado del proceso de inclusión a conveniencia. Sesgo asumido ya que no se pretende generalizar resultados a población colombiana sino comprender las relaciones existentes entre variables genéticas y no genéticas y la severidad de la infección por SARS-CoV-2.

Sesgo de información: Derivado del proceso de revisión retrospectiva (identificación de antecedentes reportados por el paciente y registrados en historia clínica) y del reporte de las encuestas realizadas. Se realizó una revisión exhaustiva de los registros clínicos del paciente por médico hospitalario y se estandarizó la encuesta inicial y de seguimiento telefónico para disminuir este sesgo.

Sesgo de clasificación: Derivado del tipo de estudio para SARS-CoV-2 se controló mediante la revisión de historias clínicas y encuestas telefónicas (usando los criterios establecidos por el Ministerio de Salud para clasificación de severidad).

Sesgo de confusión: Derivado de los factores genéticos y no genéticos. Se realizaron análisis multivariados y de sensibilidad que permitan la identificación de variables de confusión.

Sesgo de genotipificación: Derivado del proceso de genotipificación a gran escala. Se llevó a cabo procedimientos siguiendo metodologías analíticas estandarizadas y se usó una metodología de secuenciación con baja tasa de errores y se determinó la calidad de los datos de secuenciación obtenidos.

Sesgo de medición: inherente al tipo de equipos de laboratorio utilizados y a la calibración de los equipos de laboratorio utilizados. Por lo cual se utilizaron tecnologías con un margen de error mínimo en secuenciación como lo es la secuenciación por nano esferas, con apoyo de un centro de secuenciación especializado, es decir que las muestras fueron manejadas por personal entrenado y con equipos de calidad.

Sesgo de asociación: Derivado de los resultados de asociación de manifestaciones clínicas. Se llevaron a cabo análisis univariados y multivariados.

Sesgo de memoria: Derivado del proceso de encuestas realizado por los investigadores. Se estableció un límite de tiempo de 2 meses, posterior a los 30 días de inicio de sintomatología para realizar los respectivos seguimientos telefónicos, ajustados a los tiempos y horarios de cada uno de los pacientes. Adicionalmente se comparó la información telefónica reportada por el paciente con su historial clínico, disminuyendo así el sesgo de memoria e información que pudiese generar la encuesta. Cualquier persistencia sintomatológica posterior a estos 30 días de sintomatología inicial se consideró como manifestación del COVID-19 prolongado

8. PLAN DE ANÁLISIS

8.1 Técnicas de procesamiento y análisis de datos

8.1.1 Análisis genético

Se diseñó un panel *in house* de genes candidato y polimorfismos de nucleótido simple (SNPs) a partir de reportes previos de literatura científica, estudios de casos y controles, GWAS, estudios de asociación del transcriptoma por sus siglas en inglés (TWAS), cohortes prospectivas, revisiones sistemáticas y metaanálisis. El panel resultante incluyó 81 SNPs y 74 genes candidatos relacionados con la severidad de COVID-19.

Después de la obtención de sangre periférica del paciente y la extracción del ADN, se realizó un enriquecimiento de las regiones de interés (custom DNA panel) y preparación de la librería para secuenciación de siguiente generación por nanoesferas de ADN. Esta tecnología se basa en la circularización de la hebra simple de ADN mediante adaptadores en los extremos terminales después de ser desnaturalizado. Las nanoesferas se generan después mediante la amplificación de la hebra de ADN circular original como molde lo que favorece la reducción de errores de secuenciación mediante amplificación ya que emplea una ADN polimerasa de alta

fidelidad y utiliza como molde siempre la hebra original. (97) Esta secuenciación se llevó a cabo en un laboratorio externo (Gencell Pharma) usando un equipo DNBSEQ.G400 (MGI). La calidad de las lecturas se evaluó y se removieron aquellas con una calidad menor a Q30. Adicionalmente, se removieron de los análisis posteriores las variantes cuya profundidad de lectura fue inferior a 10x. El mapeo de las lecturas se realizó con el software Burrows-Wheeler Aligner (BWA) usando el genoma GRCh37/hg19 como referencia. El llamado de variantes se realizó mediante el programa bioinformático Genome Analysis Toolkit (GATK) y los VCF resultantes se analizaron mediante el programa VarSeq (Golden Helix). La profundidad y cobertura de secuenciación sobre las regiones y genes de interés se analizó utilizando la herramienta BedCov del paquete SAMtools v1.12 (98)

El análisis de variantes genéticas de los genes candidatos seleccionados se hizo aplicando los siguientes filtros de patogenicidad:

- a. Variantes reportadas previamente como patogénicas o probablemente patogénicas en la base de datos pública Clinvar.
- b. Variantes raras con $MAF \leq 0.01$ en la base de datos pública GnomAD o nuevas, cuya patogenicidad es conflictiva o desconocida se usó el predictor de patogenicidad *in silico* REVEL (99) para variantes ubicadas el marco abierto de lectura (ORF). Para los análisis posteriores se tuvieron en cuenta las variantes de ORF tipo missense con un puntaje de REVEL >0.5 dando cumplimiento al objetivo 3. Adicionalmente se realizó la clasificación de las variantes resultantes según su patogenicidad, siguiendo los lineamientos de la guía del colegio Americano de Genética Médica y Genómica usando la herramienta online (<https://franklin.genoox.com/clinical-db/home>) y se tuvieron en cuenta aquellas variantes clasificadas como patogénicas o probablemente patogénicas.

8.1.2 Análisis estadístico

Para el procesamiento de la información y análisis estadístico de los datos se usó el lenguaje R v.4.2, en el cual se implementaron diferentes paquetes estadísticos incluyendo (epiR, dplyr, viridis, lmtest, knitr, pROC, magrittr, kableExtra, ResourceSelection, boot, sjPlot, PredictABEL, rmda, dcurves, psyc). El análisis de los datos genéticos con el fin de evaluar los diferentes modelos genéticos de riesgo se llevó a cabo utilizando el programa PLINK v1.9 (<https://www.broadinstitute.org/haploview/plink>). Todos los datos se analizaron teniendo en cuenta la naturaleza de las variables, su nivel de medición y sus unidades de medida.

8.1.3 Análisis univariado

Análisis de tipo descriptivo para las variables cualitativas mediante la realización de tablas de distribución de frecuencia absoluta y relativa. Para las variables cuantitativas se empleó el test estadístico Shapiro Wilks para determinar si las variables se distribuían de manera normal y de acuerdo a este resultado, se analizaron con las medidas de tendencia central y dispersión adecuadas. De esta manera, si la variable se distribuía de manera normal se usó la media y desviación estándar, y si la variable no presentaba distribución normal, la mediana, rango, límite superior, límite inferior y rango intercuartílico. Para determinar los parámetros genético-poblacionales se calcularon las frecuencias alélicas, frecuencias genotípicas y el equilibrio de Hardy Weinberg (HWE) ($p^2 + 2pq + q^2 = 1$), el cual se estimó mediante la prueba χ^2 con 1° de libertad. Las frecuencias y la distribución de las variables incluidas en el estudio se calcularon en función de la variable dependiente severidad del COVID-19 o presencia de COVID-19 prolongado dando cumplimiento a los objetivos 1 y 2.

8.1.4 Análisis bivariado

Análisis bivariado para evaluar el comportamiento de cada variable en los pacientes severos (casos) y leves (controles), al igual que determinar la asociación entre la variable dependiente (Severidad de COVID-19) o (COVID-19 prolongado) y las variables independientes de interés. Este se llevó a cabo considerando la naturaleza de la variable dependiente cualitativa nominal, y la naturaleza de cada variable independiente a ser analizada.

Las variables incluidas en el estudio se agruparon en sociodemográficas, clínicas y genéticas. Para las variables independientes cuantitativas según su normalidad se emplearon las pruebas estadísticas U de Mann Whitney (distribución no normal) y t de Student (distribución normal). Para las variables independientes cualitativas se usó el estadístico χ^2 , valor P ($<0,05$ estadísticamente significativo) y se calculó la medida epidemiológica Odds Ratio y sus respectivos intervalos de confianza al 95%.

Para la evaluación bivariada de las variantes genéticas obtenidas de estudios de asociación previos (SNPs) se consideró el análisis bajo los siguientes modelos genéticos: alélico (D vs d), dominante (DD,Dd vs dd), recesivo (DD vs Dd,dd) y codominante (DD vs Dd vs dd), considerando D como el alelo mayor y d como el alelo de frecuencia alélica mínima. Se empleó el estadístico χ^2 con 1° de libertad (modelo dominante y recesivo) y 2° de libertad (modelo genotípico) y el Cochran-Armitage trend test (variantes moleculares que violaban el HWE) y se calculó la medida epidemiológica Odds Ratio y sus respectivos intervalos de confianza al 95% usando el paquete bioinformático PLINK. Todos los análisis fueron efectuados bajo el concepto de muestras independientes, dando cumplimiento a los objetivos 4 y 5.

8.1.5 Análisis multivariado

Considerando que la variable dependiente corresponde a un desenlace binario (severidad vs no severidad o presencia de COVID-19 prolongado vs ausencia de COVID-19 prolongado) se utilizó en modelo de regresión logística binario multivariado. Al modelo de riesgo predictivo se incorporaron las variables con significancia estadística ($p < 0.05$) derivadas del análisis bivariado y se utilizó el método de selección *Stepwise* tipo *backward* para elegir el mejor modelo basado en el criterio de información de Akaike (AIC), eligiendo aquel que presentara un menor valor de AIC.

El proceso consiste en evaluar paso a paso las variables iniciando por las variables que muestran ser más importantes para el modelo, y comparando su efecto en el logaritmo de verosimilitud con el modelo que no contiene la variable. No se sospecharon potenciales variables confusoras entre aquellas involucradas en el modelo. El test estadístico de Wald fue usado para evaluar la significancia de los coeficientes de regresión logística individuales.

Se verificó el cumplimiento de los supuestos para este modelo, no colinealidad (cálculo del VIF para variables cuantitativas y correlación para cualitativas), homocedasticidad evaluada con el test de Breusch Pagan e incorrelación de errores con el test de Durbin Watson. Se determinó el rendimiento del modelo, bondad de ajuste mediante la prueba de Hosmer-Lemeshow. Se estimó la capacidad de discriminación evaluando el área bajo la curva ROC (AUC) considerando como punto de corte para una buena discriminación un valor 0.70, dando cumplimiento al objetivo 6.

9. CONSIDERACIONES ÉTICAS:

9.1 Calificación del riesgo del estudio: Según la resolución 008430 del Ministerio de Salud de Colombia la categoría de la investigación es “investigación de riesgo mínimo”.

9.2 Método de obtención de consentimiento informado

Se realizó la obtención de este de manera presencial con cada uno de los pacientes

A cada participante del estudio se le explicaron los siguientes aspectos:

1. Que la participación en este estudio es totalmente voluntaria.
2. La naturaleza de esta investigación, su propósito, sus limitaciones, sus riesgos, sus inconvenientes, incomodidades.
3. Que se le explicaría cualquier información pertinente al resultado del estudio en caso de solicitarla.

4. Que se mantendrá absoluta confidencialidad: así los registros médicos de cada individuo permanecerían archivados por el médico tratante. Las historias médicas, los resultados de exámenes y la información proporcionada son de carácter absolutamente confidencial. Por ningún motivo se divulgaría esta información.

9.3 Métodos de minimización de riesgos del sujeto

A cada participante se le asignó un número de identificación que se usó para rotular el consentimiento informado (Anexo 1), el formulario de datos clínicos y las muestras obtenidas. La información contenida en los anexos se mantiene en las instalaciones del CIGGUR en archivadores debidamente cerrados y con acceso exclusivo de los Investigadores principales, del director del Centro de Investigación y del asistente de investigación adscrito de manera oficial al proyecto.

Todas las encuestas clínicas (Anexo 2) y consentimientos informados (Anexo 1) se escanearon y guardaron en el sistema de almacenamiento del CIGGUR con acceso exclusivo a los investigadores principales. En caso de que alguno de los dos investigadores principales se retire de la Universidad de Rosario por cualquier circunstancia, la responsabilidad será ejercida por el otro investigador principal.

En el caso que los dos investigadores tengan desafiliación a la universidad será ella quien designe el responsable de la custodia de la información relacionada con el proyecto. Los participantes se identificaron como: COV01 y la numeración se realizó de forma consecutiva. El dato de la información genómica de los pacientes se mantiene anónimo y se mantendrá en un equipo de cómputo con acceso exclusivo del Investigador principal, del director del Centro de Investigación y del asistente de investigación adscrito de manera oficial al proyecto por un tiempo no mayor a la finalización y divulgación de los resultados del estudio.

La custodia de estos datos estará hasta el final cargo de los investigadores principales y del director del centro de investigación. Se recalca que las muestras de DNA fueron obtenidas exclusivamente para el desarrollo del presente estudio cuyos productos se reflejaron en un estudio ya publicado (36) y otro que está en proceso de publicación.

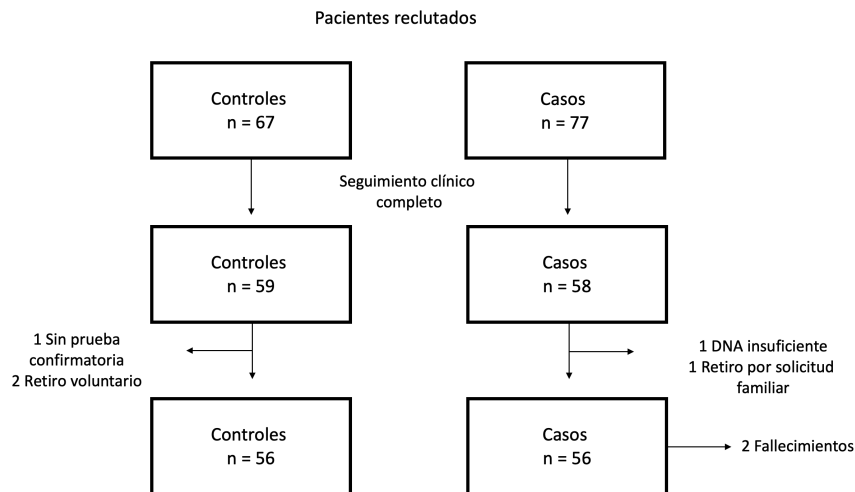
10. RESULTADOS

10.1 Definición de la muestra de estudio

El estudio incluyó 144 pacientes correspondientes a 67 controles y 77 casos, de estos 8 controles y 19 casos se excluyeron debido a pérdida de seguimiento, 2 controles y 1 caso se excluyeron dado solicitud voluntaria de retiro o solicitud familiar, 1 control se retiró por no contar con prueba confirmatoria de COVID-19 y 1 caso se excluyó dado a obtención insuficiente de DNA. Dando un total de 32 pacientes excluidos del estudio. El número de pacientes analizados fue 112 pacientes 56 controles y 56 casos, de estos, 2 pacientes fallecieron por complicaciones de COVID-19, sin embargo, el seguimiento clínico se completó con éxito gracias a la colaboración de familiares (Figura 1). En el grupo control el 80% (n=45) de los diagnósticos de COVID-19 se hizo por medio de RT-PCR, 18% (n=10) mediante prueba de antígenos y el 2%

(n=1) mediante anticuerpos. Respecto al grupo de los casos el 98% (n=55) de los diagnósticos se realizó por prueba de RT-PCR y el 2% (n=1) a través de prueba de antígenos. El 100% de las muestras obtenidas para la extracción de ADN de este estudio fueron sanguíneas. No hubo pérdidas de individuos posterior al control de calidad de los datos de secuenciación.

Figura 1. Diagrama de flujo de los participantes del estudio.



10.2 Caracterización demográfica y clínica

La muestra de pacientes estuvo compuesta en un 62.5% (n= 35) por hombres en el grupo de los casos y en un 42.9% (n= 24) en el grupo control. La mediana de la edad para los casos y los controles fue de 48 años. Las comorbilidades más frecuentes fueron hipertensión arterial y diabetes mellitus presentes en el 26.8% (n= 15) y 21.4% (n= 12) de los casos y el 14.3% (n=8) y 5.4% (n=3) de los controles, respectivamente. El índice de masa corporal tuvo una mediana para el grupo de casos de 28.4 Kg/m² y 26 Kg/m² para los controles. El 30.4% (n= 17) de los casos y el 12.5% (n= 7) de los controles presentó diagnóstico de obesidad. Adicionalmente, el 53.5% de los pacientes no tuvieron ninguna comorbilidad (n=60), mientras que el 18.9% tenía 2 o más comorbilidades, siendo la mediana de 1 en el grupo de los casos y de 0 en los controles. Con relación al antecedente de tabaquismo este estuvo presente en el 35.7% (n= 20) de los casos y en el 26.8% (n= 15) de los controles. Las demás medidas de frecuencia y tendencia central específicas para cada variable pueden verse en la Tabla 1. Los síntomas más frecuentes en ambos grupos fueron fatiga 78.6% (n =88), dolor musculo esquelético 75.9% (n =86), cefalea 67.9% (n= 76) y tos 67% (n =75). El tiempo promedio de recuperación de los síntomas fue de 23 (±12) días en el grupo de los casos y de 19 (±23) días en el grupo de los controles. (Tabla 2).

Tabla 1. Características demográficas y clínicas de la población de estudio.

Tabla 1: Características demográficas y clínicas de la población de estudio					
Variable	Controles (n = 56)	Casos (n = 56)	p-valor	ICI95%	OR
Edad	48 (59;23)†	48 (50;21)†	0.55	-	-
Sexo masculino	24 (42.8%)	35 (62.5%)	0.037*	1.04-4.74	2.22
Grupo sanguíneo					
Grupo O	39 (69.6%)	40 (71.4%)	0.836	0.48-2.46	1.09
Grupo A	13 (23.2%)	14 (25.0%)	0.825	0.46-2.62	1.10
Grupo B	3 (5.3%)	2 (3.57%)	1	0.05-5.97	0.65
Grupo AB	1 (1.78%)	0 (0.0%)	1	0.00-39.00	0
Comorbilidades					
IMC*	26 (34;19)†	28.4 (50;19)†	<0.01*	0.69-4.19	-
Obesidad	7 (12.5%)	17 (30.3%)	0.021*	1.15-8.09	3.05
Hipertensión	8 (14.2%)	15 (26.7%)	0.102	0.85-5.70	2.20
DMT2*	3 (5.3%)	12 (21.4%)	0.024*	1.28-18.16	4.82
Cáncer	3 (5.3%)	1 (1.78%)	0.618	0.03-3.19	0.32
Enfermedad coronaria	2 (3.57%)	1 (1.78%)	1	0.04-5.57	0.49
Arritmia	1 (1.78%)	1 (1.78%)	1	0.06-16.39	1
Enfermedad renal crónica	0 (0.0%)	3 (5.3%)	0.243	-	-
VIH/Inmunodeficiencia	0 (0.0%)	1 (1.78%)	1	-	-
Enfermedad autoinmune	0 (0.0%)	2 (3.57%)	0.495	-	-
Asma	2 (3.57%)	2 (3.57%)	1	0.14-7.36	1
EPOC*	0 (0.0%)	2 (3.57%)	0.495	-	-
Depresión	0 (0.0%)	1 (1.78%)	1	-	-
Epilepsia/Convulsiones	1 (1.78%)	1 (1.78%)	1	0.06-16.39	1
Otras comorbilidades	24 (42.8%)	10 (17.8%)	<0.01*	0.12-0.69	0.29
Trasplantados	1 (1.78%)	2 (3.57%)	1	0.18- 23.13	2.04
Antecedente de tabaquismo	15 (26.7%)	20 (35.7%)	0.308	0.68-3.40	1.52
Fumador activo	4 (7.14%)	3 (5.3%)	1	0.16-3.45	0.74
Número de cigarrillos por día	0 (40;0)	0 (70;0)	0.72	-	0.72
Años de tabaquismo	0 (30;0)†	0 (40;0)†	0.55	-	-
IPA*	0 (34;0)†	0 (52;0)†	0.75	-	-
Uso de medicamentos	35	31	0.44	0.35-1.58	0.74
Uso crónico de esteroides	1	2	1	0.18-23.13	2.04
Número de comorbilidades	0 (4;0)†	1 (5;0)†	<0.01*	-	-
Ninguna comorbilidad	39 (69.6%)	21 (37.5%)	<0.01*	0.12-0.57	0.26
Inversa Ninguna comorbilidad	17 (30.3%)	35 (62.5%)	<0.01*	1.74-8.39	3.82
2 o más comorbilidades	7 (12.2%)	15 (26.7%)	0.057*	0.95-6.88	2.56

*Estadísticamente significativo, p-valor < 0.05; EPOC, Enfermedad pulmonar obstructiva crónica; IMC, índice de masa corporal; IPA, índice paquete año de cigarrillos; IC, intervalos de confianza; OR, Razón de Odds

† Variable de distribución no normal, se presenta su mediana (Ls límite superior; Li Límite inferior)

Tabla 2. Síntomas de COVID-19 de la población de estudio

Tabla 2: Síntomas de COVID-19 de la población de estudio.					
Variable	Síntomas		p-valor	ICI95%*	OR*
	Controles (n = 56)	Casos (n = 56)			
Disnea	17 (30.3%)	49 (87.5%)	<0.01*	6.05-42.60	16.06
Tos	28 (50.0%)	47 (83.9%)	<0.01*	2.16-12.65	5.22
Hemoptisis	2 (3.57%)	7 (12.5%)	0.162	0.76-19.46	3.86
Anosmia	35 (63.5%)	18 (32.1%)	<0.01*	0.13-0.62	0.28
Ageusia	29 (51.7%)	24 (42.8%)	0.344	0.33-1.47	0.70
Fiebre	22 (29.2%)	42 (75.0%)	<0.01*	2.07-10.41	4.64
Odinofagia	15 (26.7%)	27 (48.2%)	0.019*	1.15-5.61	2.54
Fatiga	36 (64.2%)	52 (92.8%)	<0.01*	2.28-22.91	7.22
Dolor musculoesquelético	42 (75.0%)	43 (76.7%)	0.825	0.46-2.62	1.10
Diarrea	12 (21.4%)	20 (35.7%)	0.094	0.88-4.72	2.04
Nauseas/emesis	10 (17.8%)	12 (21.4%)	0.634	0.49-3.20	1.25
Alt_estado mental	8 (14.2%)	15 (26.7%)	0.102	0.85-5.70	2.20
Congestión nasal	21 (37.5%)	22 (39.2%)	0.846	0.50-2.31	1.08
Cefalea	41 (73.2%)	35 (62.5%)	0.225	0.27-1.36	0.61
Tiempo de recuperación	19.2 (±23.2)	22.6 (±12.2)	<0.01*	-	-
Secuelas	22 (39.2%)	44 (78.5%)	<0.01*	2.46-13.04	5.67

*Estadísticamente significativo, p-valor < 0.05; IC, intervalos de confianza; OR, Razón de Odds.

Respecto a la presencia de COVID-19 persistente, este se presentó en el 78.6% (n =44) de los casos y en el 39.3% (n = 22) de los controles. Los síntomas de COVID-19 prolongado más frecuentes en ambos grupos fueron “signos y síntomas comunes” presentes en el 41% (n= 46) y “signos y síntomas neurológicos” en el 33.9% (n = 38) (Tabla 3). La caracterización de los pacientes por la presencia o ausencia de COVID-19 prolongado puede verse en la Tabla 4. Se observa una mediana de edad de 45 años para los pacientes sin cuadros prolongados y de 48 años para los pacientes con cuadros prolongados, el sexo femenino predominó en los pacientes con COVID-19 prolongado 51.6% (n = 34), mientras que el masculino fue mayor en pacientes sin COVID-19 prolongado 73.9% (n = 34). Se observa la distribución de la prevalencia de los síntomas en pacientes con y sin COVID-19 prolongado en la tabla 5.

Tabla 3. Signos y síntomas de COVID-19 prolongado en la población de estudio

Tabla 3: Signos y síntomas de COVID19 prolongado de la población de estudio					
Signos y síntomas de COVID19 prolongado					
Variable	Controles (n = 56)	Casos (n = 56)	p-valor	ICI95%*	OR*
Depresión	2 (3.57%)	22 (39.2%)	<0.01*	3.79-77.64	17.15
Cansancio	11 (19.6%)	27 (48.2%)	<0.01*	1.64-8.84	3.81
Confusión	2 (3.57%)	19 (33.9%)	<0.01*	3.04-73.13	13.86
Vértigo	0 (0.0%)	8 (14.2%)	<0.01*	-	-
Persistencia anosmia/ageusia	7 (12.5%)	6 (10.7%)	0.9717	026- 2.68	0.84
Signos y síntomas comunes	12 (21.4%)	34 (60.7%)	<0.01*	2.46-13.04	5.67
Signos y síntomas neurológicos	10 (17.8%)	28 (50.0%)	<0.01*	1.94-10.8	4.6
Signos y síntomas psiquiátricos	2 (3.57%)	20 (35.7%)	<0.01*	3.24-137.49	14.68
Signos y síntomas musculoesqueléticos	3 (5.35%)	14 (25.0%)	<0.01*	1.59-21.8	5.89
Signos y síntomas respiratorios	6 (10.7%)	16 (28.5%)	<0.01*	1.19-9.3	3.3
Signos y síntomas cardiovasculares	0 (0.0%)	10 (17.8%)	<0.01*	-	-
Signos y síntomas "otros"	0 (0.0%)	2 (3.57%)	0.495	-	-
Tiempo de persistencia	0 (210;0)†	55 (350;0)†	<0.01*	-	-

*Estadísticamente significativo, p-valor < 0.05; IC, intervalos de confianza;OR, Razón de Odds

† Variable de distribución no normal se presenta su mediana (Ls límite superior; Li Límite inferior)

10.3 Análisis de variables clínicas y demográficas:

Los análisis de los datos clínicos muestran múltiples asociaciones estadísticamente significativas con las variables de desenlace. Iniciando por severidad, entre las variables demográficas y clínicas asociadas positivamente se encuentran el sexo masculino ($p = 0.037$; OR 2.22; 95% IC = 1.04 – 4.74), obesidad ($p = 0.021$; OR 3.05; 95% IC = 1.15 – 8.09), diabetes mellitus tipo 2 ($p = 0.021$; OR 4.82; 95% IC = 1.28 – 18.16) y número de comorbilidades ($p=0.001$). Por otra parte, asociadas negativamente se encuentran “otras comorbilidades” ($p = 0.004$, OR 0.29; 95% IC = 0.12 – 0.69) y la presencia de ninguna comorbilidad ($p = 0.00127$, OR 0.26, 95% CI = 0.12 – 0.57). Información acerca de las variables demográficas y antecedentes clínicos asociadas con severidad de COVID-19 se presentan en la Tabla 1.

Diferentes síntomas mostraron asociación estadísticamente significativa con severidad. Entre ellas encontramos: síntomas respiratorios tales como disnea ($p < 0.001$; OR 16.06; 95% IC = 6.05 – 42.60), tos ($p < 0.001$; OR 5.22; 95% IC = 2.16 – 12.65) y odinofagia ($p = 0.019$; OR 2.54; 95% IC = 1.15 – 5.61); síntomas sistémicos incluyendo fiebre ($p < 0.001$; OR 4.64; 95% IC = 2.07 – 10.41) y fatiga ($p < 0.001$; OR 7.22; 95% IC = 2.28 – 22.91). El tiempo de recuperación mostró significancia estadística ($p < 0.001$), al igual que la presencia de secuelas ($p < 0.001$; OR 5.67; 95% IC = 2.46 – 13.04). Cabe resaltar anosmia ($p = 0.001$; OR 0.28; 95% IC = 0.28) como el único síntoma que mostró una asociación negativa. Sin embargo, debe tenerse en cuenta su mayor frecuencia en el grupo de controles. Información acerca de signos y síntomas con severidad de COVID-19 se presentan en la Tabla 2.

En la tabla 3 observamos los diferentes signos y síntomas de COVID-19 prolongado en casos y controles. Todas las variables a excepción de Anosmia/Ageusia y “Signos y síntomas otros” mostraron significancia estadística con la severidad del cuadro. Se resalta la aparición de depresión ($p < 0.001$; OR 17.15; 95% IC = 3.79 – 77.64), signos y síntomas psiquiátricos ($p < 0.001$; OR 14.68; 95% IC = 3.24 – 137.49), confusión ($p < 0.001$; OR 13.86; 95% IC = 3.04 – 73.13), “signos y síntomas comunes” ($p < 0.001$; OR 5.67; 95% IC = 2.46 – 13.04) y el tiempo de persistencia de estos ($p < 0.001$) en el grupo de casos.

Por otra parte, al evaluar los resultados generados tomando como variable dependiente la presencia de secuelas, no se observó ninguna asociación estadísticamente significativa entre este desenlace y las diferentes características demográficas y antecedentes clínicos evaluados (Tabla 4). Sin embargo, se observaron múltiples síntomas durante la enfermedad asociados con la aparición de COVID-19 persistente, entre ellos disnea ($p < 0.001$; OR 3.52; 95% IC = 1.59 – 7.77), tos ($p = 0.010$, OR 3.12, 95% IC = 1.38 – 7.05), ageusia ($p = 0.026$; OR 2.39; 95% IC = 1.10 – 5.21) y fiebre ($p < 0.001$; OR 4.94; 95% IC = 1.84 – 13.24). Información adicional acerca de estas variables puede ser consultada en la tabla 5.

Tabla 4. Características clínicas y demográficas de los pacientes con y sin COVID-19 prolongado.

Variable	Sin COVID-19 prolongado (n = 46)	Con COVID-19 prolongado (n = 66)	p-valor	IC95%*	OR*
Edad	45 (60;23)†	48 (60;21)†	0.26	-	-
Sexo masculino	34 (73.9%)	32(48.4%)	0.287	0.31-1.42	0.66
Grupo sanguíneo					
Grupo O	33 (71.7%)	46 (66.6%)	0.81	0.40-2.08	0.91
Grupo A	11 (23.9%)	16 (24.2%)	0.96	0.42-2.46	1.02
Grupo B	1 (2.17%)	4 (6.06%)	0.647	0.27-145.99	2.87
Grupo AB	1 (2.17%)	0 (0.0%)	0.4107	0.00-27.0	-
Comorbilidades					
IMC*	26 (44;19)†	28 (50;20) †	0.13	-	-
Obesidad	8 (17.3%)	16 (24.2%)	0.385	0.59-3.92	1.52
Hipertensión	7 (15.2%)	16 (24.2%)	0.245	0.67-4.76	1.78
DMT2*	5 (10.8%)	10 (15.15%)	0.583	0.47-4.61	1.46
Cáncer	1 (2.17%)	3 (4.54%)	0.643	0.22-21.27	2.14
Enfermedad coronaria	2 (4.34%)	1 (1.51%)	0.567	0.03-3.85	0.34
Arritmia	1 (2.17%)	1 (1.51%)	1	0.04-11.36	0.69
Enfermedad renal crónica	1 (2.17%)	2 (3.03%)	1	0.12-15.98	1.41
VIH/Inmunodeficiencia	1 (2.17%)	0 (0.0%)	0.411	-	-
Enfermedad autoinmune	0 (0.0%)	2 (3.03%)	0.512	-	-
Asma	1 (2.17%)	3 (4.54%)	0.643	0.22-21.27	-
EPOC*	0 (0.0%)	2 (3.03%)	0.512	-	-
Depresión	0 (0.0%)	1 (1.51%)	1	-	-
Epilepsia/Convulsiones	2 (4.34%)	0 (0.0%)	0.167	-	-
Otras comorbilidades	17 (36.9%)	20 (30.3%)	0.461	0.33-1.64	0.74
Trasplantados	1 (2.17%)	2 (3.03%)	1	0.12-15.98	1.41
Antecedente de tabaquismo	11 (23.9%)	24 (36.3%)	0.162	0.78-4.22	1.82
Fumador activo	1 (2.17%)	6 (9.09%)	0.237	0.52-38.71	4.50
Número de cigarrillos por día	0 (40;0)†	0 (70;0)†	0.29	-	-
Años de tabaquismo	0 (30;0)†	0 (40;0)†	0.12	-	-
IPA*	0 (34;0)†	0 (52;0)†	0.27	-	-
Uso de medicamentos	27 (58.6%)	39 (59.0%)	0.967	0.47-2.18	1.02
Uso crónico de esteroides	0 (0.0%)	3 (4.54%)	0.267	-	-
Número de comorbilidades	0 (4;0)†	0 (5;0)†	0.29	-	-
Ninguna comorbilidad	27 (58.6%)	33 (50.0%)	0.364	0.33-1.50	0.70
Inversa Ninguna comorbilidad	19 (41.3%)	33 (50.0%)	0.364	0.66-3.04	1.42
2 o más comorbilidades	7 (15.2%)	15 (22.7%)	0.32	0.61-4.41	1.64
Severidad	22 (47.8%)	44 (66.6%)	<0.01*	2.46-13.04	5.67

*Estadísticamente significativo, p-valor < 0.05; EPOC, Enfermedad pulmonar obstructiva crónica; IMC, índice de masa corporal; IPA, índice paquete año de cigarrillos; IC, intervalos de confianza;OR, Razón de Odds

† Variable de distribución no normal, se presenta su mediana (Ls límite superior; Li Límite inferior)

Tabla 5. Síntomas de COVID-19 en pacientes con y sin COVID-19 prolongado.

Variable	Síntomas		p-valor	IC95%*	OR*
	Sin COVID-19 prolongado (n = 46)	Con COVID-19 prolongado (n = 66)			
Disnea	19 (41.3%)	47 (71.2%)	<0.01*	1.59-7.77	3.52
Tos	24 (52.1%)	51 (77.2%)	<0.01*	1.38-7.05	3.12
Hemoptisis	2 (4.34%)	7 (10.6%)	0.304	0.52-13.18	2.61
Anosmia	20 (43.4%)	33 (50%)	0.496	0.61-2.77	1.30
Ageusia	16 (34.7%)	37 (56.0%)	0.026*	1.10-5.21	2.39
Fiebre	19 (41.3%)	45 (68.1%)	<0.01*	1.39-6.66	3.05
Odinofagia	14 (30.4%)	28 (42.2%)	0.197	0.76-3.73	1.68
Fatiga	29 (63.%)	59 (89.3%)	<0.01*	1.84-13.24	4.94
Dolor musculoesquelético	32 (69.5%)	53 (80.3%)	0.191	0.74-4.27	1.78
Diarrea	9 (19.5%)	23 (34.%)	0.078	0.91-5.34	2.20
Nauseas/emesis	8 (17.3%)	14 (21.2%)	0.617	0.49-3.35	1.28
Alteración del estado de consciencia	7 (15.2%)	16 (24.2%)	0.245	0.67-4.76	1.78
Congestion nasal	15 (32.6%)	28 (42.2%)	0.293	0.69-3.34	1.52
Cefalea	31 (67.3%)	45 (68.1%)	0.930	0.46-2.32	1.04
Tiempo recuperación	10.7 (±8.4)	28.0 (±20.1)	<0.01*	-	-

*Estadísticamente significativo, p-valor < 0.05;IC, intervalos de confianza;OR, Razón de Odds

10.4 Análisis de calidad de los datos genéticos

Se evaluó la profundidad y cobertura alcanzada tanto para la secuenciación de los SNP como para la secuenciación de los genes candidatos. Para los SNPs se eliminaron aquellos con una profundidad menor del 10X en más del 5% de los casos, estos correspondieron a rs143334143, rs7310667, rs11085727, rs13050728, rs622568, rs1981555 y rs113661667 (n=7). Los genotipos de SNPs en pacientes con profundidades menores de 10X fueron designados como desconocidos (./.) y no fueron incluidos en el análisis de asociación para ese polimorfismo en particular. La profundidad promedio de los SNP fue de 832.3x. (Tabla suplementaria 1). Con relación a los genes candidato para el análisis de cobertura se utilizaron los transcritos principales según ENSEMBL (<https://www.ensembl.org/index.html>) tomado como regiones blanco las secuencias codificantes y +/- 50 pb downstream o upstream de la región de empalme. La cobertura de las regiones de interés (179 kbp) fue mayor de 99% para una profundidad de 20X en todos los pacientes. (Tabla suplementaria 2). La profundidad promedio para los genes de interés fue de 1206X, con un rango entre 503X y 1988X (Tabla suplementaria 3)

10.5 Análisis de polimorfismo de nucleótido simple (SNP)

10.5.1 Parámetros genético poblacionales

Los parámetros genéticos poblacionales incluyendo frecuencias alélicas, genotípicas y HWE para severidad y presencia de COVID-19 prologando se presentan en las tablas 6 y 7, respectivamente. Todos los polimorfismos se encontraron en equilibrio de Hardy-Weinberg a excepción de los siguientes: rs41264915, rs2232354, rs147509469, rs4424872, rs73510898 y rs368565.

Tabla 6. Frecuencias alélicas y genotípicas de casos y controles de severidad de COVID-19.

Tabla 6: Frecuencias alélicas y genotípicas para los casos y controles												
SNP	Gen	Frecuencias alélicas controles		Frecuencias alélicas en los casos		Genotipo Controles			Genotipo Casos			HWE
		WT	Alt	WT	Alt	WT/WT	WT/Alt	Alt/Alt	WT/WT	WT/Alt	Alt/Alt	
rs114301457	<i>EFNA4</i>	1	0.0	0.99	0.01	1	0.0	0.0	0.98	0.02	0.0	1
rs7528026	<i>TRIM46</i>	1	0.0	0.96	0.04	1	0.0	0.0	0.91	0.09	0.0	1
rs41264915	<i>THBS3</i>	0.97	0.03	0.98	0.02	0.96	0.02	0.02	0.98	0.0	0.02	<0.01*
rs1123573	<i>BCL11A</i>	0.79	0.21	0.76	0.24	0.62	0.34	0.04	0.54	0.45	0.02	0.27
rs2232354	<i>IL1RN</i>	0.86	0.14	0.72	0.28	0.04	0.21	0.75	0.14	0.27	0.59	<0.01*
rs147509469	†	0.97	0.03	1	0.0	0.96	0.02	0.02	1	0.0	0.0	<0.01*
rs73062389	<i>SLC6A20</i>	0.95	0.05	0.95	0.05	0.91	0.07	0.02	0.89	0.11	0.0	0.27
rs2271616	<i>SLC6A20</i>	0.87	0.13	0.89	0.11	0.75	0.23	0.02	0.82	0.14	0.04	0.19
rs2531743	<i>SLC6A20</i>	0.71	0.29	0.77	0.23	0.05	0.48	0.46	0.02	0.43	0.55	0.89
rs72893671	†	0.96	0.04	0.91	0.09	0.91	0.09	0.0	0.82	0.18	0.0	1
rs17713054	<i>LOC107986083</i>	0.99	0.01	0.92	0.08	0.98	0.02	0.0	0.84	0.16	0.0	1
rs17713054	<i>LOC107986083</i>	0.99	0.01	0.92	0.08	0.98	0.02	0.0	0.84	0.16	0.0	1
rs71325088	<i>LOC107986083</i>	0.99	0.01	0.92	0.08	0.98	0.02	0.0	0.84	0.16	0.0	1
rs10490770	<i>LZTFL1</i>	0.99	0.01	0.92	0.08	0.98	0.02	0.0	0.84	0.16	0.	1
rs11385942	<i>LZTFL1</i>	0.99	0.01	0.91	0.09	0.98	0.02	0.0	0.82	0.18	0.0	1
rs35081325	<i>LZTFL1</i>	0.99	0.01	0.92	0.08	0.98	0.02	0.0	0.84	0.16	0.0	1
rs73064425	<i>LZTFL1</i>	0.99	0.01	0.92	0.08	0.98	0.02	0.0	0.84	0.16	0.0	1
rs71325091	<i>CCR9</i>	0.96	0.04	0.92	0.08	0.93	0.07	0.0	0.84	0.16	0.0	1
rs13433997	†	0.92	0.08	0.88	0.12	0.84	0.16	0.0	0.77	0.23	0.0	0.6
rs34438204	<i>XCR1</i>	0.96	0.04	0.92	0.08	0.91	0.09	0.0	0.84	0.16	0.0	1
rs7642320	<i>XCR1</i>	0.88	0.12	0.86	0.14	0.77	0.23	0.0	0.71	0.29	0.0	0.21
rs9877748	<i>XCR1</i>	0.91	0.09	0.88	0.12	0.84	0.14	0.02	0.77	0.23	0.0	1
rs13069742	<i>XCR1</i>	0.91	0.09	0.88	0.12	0.84	0.14	0.02	0.77	0.23	0.0	1
rs35110864	†	0.96	0.04	0.92	0.08	0.91	0.09	0.0	0.84	0.16	0.0	1
rs13085367	<i>LOC105377067</i>	0.96	0.04	0.92	0.08	0.93	0.07	0.0	0.86	0.12	0.02	0.31

Tabla 6: Frecuencias alélicas y genotípicas para los casos y controles

SNP	Gen	Frecuencias alélicas controles		Frecuencias alélicas en los casos		Genotipo Controles			Genotipo Casos			HWE
rs4443214	<i>LOC105377067</i>	0.94	0.06	0.91	0.09	0.88	0.12	0.0	0.84	0.14	0.03	0.48
rs35775079	<i>CCR3</i>	0.99	0.01	0.96	0.04	0.98	0.02	0.0	0.91	0.09	0.0	1
rs11919389	†	0.79	0.21	0.81	0.19	0.61	0.36	0.04	0.66	0.3	0.04	1
rs343320	<i>PLSCR1</i>	0.91	0.09	0.96	0.04	0.84	0.15	0.02	0.93	0.07	0.0	0.35
rs56162149	<i>ACSL6</i>	0.83	0.17	0.79	0.21	0.68	0.3	0.02	0.64	0.29	0.07	0.55
rs9271609	†	0.66	0.34	0.65	0.35	0.36	0.61	0.04	0.46	0.38	0.16	0.41
rs2496644	<i>LINC01276</i>	0.26	0.74	0.27	0.73	0.05	0.41	0.54	0.07	0.39	0.54	0.81
rs1886814	<i>FOXP4-AS1</i>	0.77	0.23	0.77	0.23	0.59	0.36	0.05	0.59	0.36	0.05	1
rs28368148	<i>IFNA10</i>	0.99	0.01	0.99	0.01	0.98	0.02	0.0	0.98	0.02	0.0	1
rs1834952085, rs505922	<i>ABO</i>	0.84	0.16	0.82	0.18	0.7	0.29	0.02	0.68	0.29	0.04	1
rs529565	<i>ABO</i>	0.84	0.16	0.78	0.22	0.7	0.29	0.02	0.61	0.34	0.05	1
rs61882275	<i>ELF5</i>	0.61	0.39	0.54	0.46	0.32	0.57	0.11	0.32	0.45	0.23	0.7
rs10774671	<i>OAS1</i>	0.24	0.76	0.18	0.82	0.05	0.38	0.57	0.07	0.21	0.71	0.25
rs2660	<i>OAS1</i>	0.18	0.82	0.15	0.85	0.02	0.32	0.66	0.07	0.16	0.77	0.18
rs10850097	<i>OAS1</i>	0.21	0.79	0.16	0.84	0.04	0.36	0.61	0.07	0.18	0.75	0.21
rs6489867	<i>OAS1</i>	0.23	0.77	0.18	0.82	0.04	0.39	0.57	0.09	0.18	0.73	0.24
rs7955267	<i>OAS3</i>	0.21	0.79	0.2	0.8	0.02	0.39	0.59	0.11	0.18	0.71	0.24
rs56106917	<i>FBRSL1</i>	0.74	0.26	0.71	0.29	0.57	0.34	0.09	0.5	0.41	0.09	0.48
rs9577175	†	0.79	0.21	0.72	0.28	0.59	0.39	0.02	0.5	0.45	0.05	0.21
rs4424872	<i>RGMA</i>	0.04	0.96	0.03	0.97	0.04	0.02	0.95	0.02	0.02	0.96	<0.01*
rs117169628	<i>SLC22A31</i>	0.88	0.12	0.87	0.13	0.77	0.23	0.0	0.79	0.16	0.05	0.47
rs79600142	<i>CRHR1</i>	0.87	0.13	0.94	0.06	0.73	0.27	0.0	0.89	0.09	0.02	1
rs62054835	<i>MAPT-AS1</i>	0.87	0.13	0.95	0.05	0.73	0.27	0.0	0.91	0.07	0.02	1
rs112572874	<i>MAPT</i>	0.86	0.14	0.94	0.06	0.71	0.29	0.0	0.89	0.09	0.02	1
rs1819040	<i>KANSL1</i>	0.85	0.15	0.94	0.06	0.73	0.23	0.04	0.89	0.09	0.02	0.11
rs2532300	<i>KANSL1</i>	0.87	0.13	0.94	0.06	0.73	0.27	0.0	0.89	0.09	0.02	1
rs3848456	†	0.75	0.25	0.77	0.23	0.52	0.46	0.02	0.57	0.39	0.04	0.12
rs77534576	†	0.83	0.17	0.85	0.15	0.66	0.34	0.0	0.73	0.23	0.04	0.73

Tabla 6: Frecuencias alélicas y genotípicas para los casos y controles

SNP	Gen	Frecuencias alélicas controles		Frecuencias alélicas en los casos		Genotipo Controles			Genotipo Casos			HWE
rs12610495	<i>DPP9</i>	0.78	0.22	0.71	0.29	0.61	0.34	0.05	0.55	0.32	0.12	0.21
rs2109069	<i>DPP9</i>	0.75	0.25	0.70	0.30	0.55	0.39	0.05	0.52	0.36	0.12	0.48
rs2277732	<i>DPP9</i>	0.77	0.23	0.71	0.29	0.59	0.36	0.05	0.54	0.34	0.12	0.33
rs4804803	<i>CD209</i>	0.82	0.18	0.85	0.15	0.68	0.29	0.04	0.75	0.20	0.05	0.18
rs73510898	<i>ZGLP1</i>	0.96	0.04	0.96	0.04	0.91	0.09	0.0	0.95	0.02	0.04	<0.01*
rs74956615	<i>FDX2</i>	0.94	0.06	0.98	0.02	0.88	0.12	0.0	0.96	0.04	0.0	1
rs34536443	<i>TYK2</i>	0.96	0.04	0.98	0.02	0.93	0.07	0.0	0.96	0.04	0.0	1
rs429358	<i>APOE</i>	0.88	0.12	0.95	0.05	0.79	0.18	0.04	0.89	0.11	0.0	0.2
rs368565	<i>FUT2</i>	0.34	0.66	0.49	0.51	0.16	0.36	0.48	0.35	0.29	0.36	<0.01*
rs4801778	<i>PLEKHA4</i>	0.85	0.15	0.88	0.12	0.73	0.23	0.04	0.77	0.21	0.02	0.44
rs17860115	<i>IFNAR2</i>	0.52	0.48	0.48	0.52	0.25	0.54	0.21	0.25	0.46	0.29	1
rs2300370	<i>IFNAR2</i>	0.54	0.46	0.46	0.54	0.27	0.54	0.2	0.23	0.46	0.3	1
rs2252639	<i>IFNAR2</i>	0.52	0.48	0.53	0.47	0.23	0.57	0.2	0.29	0.48	0.23	0.7
rs2236757	<i>IFNAR2</i>	0.46	0.54	0.48	0.52	0.18	0.57	0.25	0.27	0.43	0.30	0.1
rs2300371	<i>IFNAR2</i>	0.54	0.46	0.55	0.45	0.25	0.57	0.18	0.34	0.43	0.23	1
rs8178521	<i>IL10RB</i>	0.76	0.24	0.74	0.26	0.57	0.38	0.05	0.61	0.27	0.12	0.13
rs147418963	<i>LINC00649</i>	0.83	0.17	0.79	0.21	0.66	0.34	0.0	0.62	0.34	0.02	0.35
rs2298660	<i>TMPRSS2</i>	0.79	0.21	0.81	0.19	0.62	0.34	0.04	0.68	0.27	0.05	0.76
rs2298661	<i>TMPRSS2</i>	0.78	0.22	0.82	0.18	0.61	0.34	0.05	0.7	0.25	0.05	0.38
rs3787946	<i>TMPRSS2</i>	0.78	0.22	0.80	0.20	0.61	0.34	0.05	0.66	0.29	0.05	0.57

SNP (polimorfismo de nucleótido simple); WT(Wild Type); Alt (alterno); HWE (Equilibrio de Hardy Weinberg)

† Polimorfismo no se encuentra en un gen específico reportado en la literatura actual.

* Polimorfismo que viola el HWE para todos los sujetos

SNP (polimorfismo de nucleótido simple); WT (Wild Type); Alt (alterno); HWE (Equilibrio de Hardy Weinberg)

† Polimorfismo no se encuentra en un gen específico reportado en la literatura actual.

* Polimorfismo que viola el HWE para todos los sujetos

10.5.2. Análisis Bivariado con severidad de COVID-19

El análisis bivariado entre los polimorfismos genéticos con la severidad del COVID-19 reveló asociación estadísticamente significativa para 13 de ellos (rs7528026, rs2232354, rs17713054, rs71325088, rs10490770, rs11385942, rs35081325, rs73064425, rs35775079, rs62054835, rs112572874, rs1819040, rs368565). Dentro de ellos se destacan las siguientes variantes, las cuales obtuvieron la mayor fuerza de asociación: rs11385942 ($p < 0.01$; OR=10.88; 95% CI=1.36-86.51), rs10490770 ($p < 0.01$; OR=9.699; 95% CI=1.20-77.89), rs35081325 ($p < 0.01$; OR=9.699; 95% CI=1.20-77.89) y rs73064425 ($p < 0.01$; OR=9.699; 95% CI=1.20-77.89) en gen *LZTFL1*, la variante rs71325088 en el gen LOC107986083 ($p < 0.01$; OR=9.699; 95% CI=1.20-77.89) y la variante rs35775079 en el gen *CCR3* ($p = 0.01$; OR=8.538; 95% CI=1.05-69.45) (Tabla 8). Todas las variantes presentaron asociación estadísticamente significativa bajo los modelos alélicos y el Cochran-Armitage test.

Tabla 8. Análisis bivariado entre los polimorfismos genéticos con la severidad del COVID-19.

SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico en casos	Genotipo alélico en controles	X2	df	Valor p	OR	IC 95%
rs114301457	C/T	C	Test de Genotipo (2df)	0/1/55	0/0/56	-	-	-	-	-
			Test de Cochran armitage	1/111	0/112	1.009	1	0.3151	-	-
			Alélico	1/111	0/112	1.004	1	0.3162	-	-
			Dominante	1/55.	0/56	-	-	-	-	-
			Recesivo	0/56	0/56	-	-	-	-	-
rs7528026	G/A	G	Test de Genotipo (2df)	0/5/51	0/0/56	-	-	-	-	-
			Test de Cochran armitage	5/107	0/112	5.234	1	0.02215*	-	-
			Alélico	5/107	0/112	5.114	1	0.02373*	-	-
			Dominante	5/51.	0/56	-	-	-	-	-
			Recesivo	0/56	0/56	-	-	-	-	-
rs41264915	A/G	A	Test de Genotipo (2df)	0/5/51	0/6/50	-	-	-	-	-
			Test de Cochran armitage	5/107	6/106	0.1008	1	0.7509	-	-
			Alélico	5/107	6/106	0.0956	1	0.7572	0.8255	0.2445-
			Dominante	05/51.	06/50.	-	-	-	-	-
			Recesivo	0/56	0/56	-	-	-	-	-
rs1123573	A/G	A	Test de Genotipo (2df)	1/25/30	2/19/35	-	-	-	-	-
			Test de Cochran armitage	27/85	23/89	0.4751	1	0.4907	-	-
			Alélico	27/85	23/89	0.412	1	0.521	1.229	0.6542-
			Dominante	26/30	21/35	-	-	-	-	-
			Recesivo	01/55.	02/54.	-	-	-	-	-

Tabla 8: Análisis de asociación genético para COVID-19 severo

SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico en casos	Genotipo alélico en controles	X2	df	Valor p	OR	IC 95%
rs2232354	T/G	T	Test de Genotipo (2df)	8/15/33	2/12/42	-	-	-	-	-
			Test de Cochran armitage	31/81	16/96	4.759	1	0.02914*	-	-
			Alélico	31/81	16/96	6.058	1	0.01384*	2.296	1.173-4.496
			Dominante	23/33	14/42	-	-	-	-	-
			Recesivo	08/48.	02/54.	-	-	-	-	-
rs147509469	G/A	A	Test de Genotipo (2df)	0/0/56	1/01/54	-	-	-	-	-
			Test de Cochran armitage	0/112	3/109	1.829	1	0.1762	-	-
			Alélico	0/112	3/109	3.041	1	0.0812	0	-
			Dominante	0/56	02/54.	-	-	-	-	-
			Recesivo	0/56	01/55.	-	-	-	-	-
rs73062389	A/G	G	Test de Genotipo (2df)	0/6/50	1/04/51	-	-	-	-	-
			Test de Cochran armitage	6/106	6/106	0	1	1	-	-
			Alélico	6/106	6/106	0	1	1	1	0.3125-3.2
			Dominante	06/50.	05/51.	-	-	-	-	-
			Recesivo	0/56	01/55.	-	-	-	-	-
rs2271616	G/T	T	Test de Genotipo (2df)	2/08/46	1/13/42	-	-	-	-	-
			Test de Cochran armitage	12/100	15/97	0.3397	1	0.56	-	-
			Alélico	12/100	15/97	0.379	1	0.5381	0.776	0.3456-1.742
			Dominante	10/56.	14/42	-	-	-	-	-
			Recesivo	02/54.	01/55.	-	-	-	-	-
rs2531743	G/A	A	Test de Genotipo (2df)	1/24/31	3/27/26	-	-	-	-	-
			Test de Cochran armitage	26/86	33/79	1.364	1	0.2428	-	-
			Alélico	26/86	33/79	1.127	1	0.2883	0.7237	0.398-1.316
			Dominante	25/31	30/26	-	-	-	-	-
			Recesivo	01/55.	03/53.	-	-	-	-	-
rs72893671	T/A	T	Test de Genotipo (2df)	0/10/46	0/5/51	-	-	-	-	-
			Test de Cochran armitage	10/102	5/107	1.924	1	0.1654	-	-
			Alélico	10/102	5/107	1.786	1	0.1814	2.098	0.5649-0.6933
			Dominante	10/46.	05/51.	-	-	-	-	-
			Recesivo	0/56	0/56	-	-	-	-	-
rs17713054	G/A	G	Test de Genotipo (2df)	0/9/47	0/1/55	-	-	-	-	-
			Test de Cochran armitage	9/103	1/111	7.027	1	0.008027*	-	-
			Alélico	9/103	1/111	6.699	1	6*	9.699	1.208-77.89
			Dominante	09/47.	01/55.	-	-	-	-	-
			Recesivo	0/56	0/56	-	-	-	-	-
rs71325088	T/C	T	Test de Genotipo (2df)	0/9/47	0/1/55	-	-	-	-	-
			Test de Cochran armitage	9/103	1/111	7.027	1	0.008027*	-	-
			Alélico	9/103	1/111	6.699	1	6*	9.699	1.208-77.89
			Dominante	09/47.	01/55.	-	-	-	-	-
			Recesivo	0/56	0/56	-	-	-	-	-
rs10490770	T/C	T	Test de Genotipo (2df)	0/9/47	0/1/55	-	-	-	-	-
			Test de Cochran armitage	9/103	1/111	7.027	1	0.008027*	-	-
			Alélico	9/103	1/111	6.699	1	6*	9.699	1.208-77.89
			Dominante	09/47.	01/55.	-	-	-	-	-
			Recesivo	0/56	0/56	-	-	-	-	-
rs11385942	D/A	D	Test de Genotipo (2df)	0/10/46	0/1/55	-	-	-	-	-
			Test de Cochran armitage	10/102	1/111	8.166	1	0.004269*	-	-
			Alélico	10/102	1/111	7.744	1	9*	10.88	1.369-86.51
			Dominante	10/46.	01/55.	-	-	-	-	-
			Recesivo	0/56	0/56	-	-	-	-	-

Tabla 8: Análisis de asociación genético para COVID-19 severo

SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico en casos	Genotipo alélico en controles	X2	df	Valor p	OR	IC 95%
rs7642320	A/G	A	Test de Genotipo (2df)	0/16/40	0/13/43	-	-	-	-	-
			Test de Cochran armitage	16/96	13/99	0.4188	1	0.5175	-	-
			Alélico Dominante	16/96	13/99	0.3565	1	0.5505	1.269	0.5796-2.779
			Recesivo	16/40	13/43	-	-	-	-	-
rs9877748	A/G	A	Test de Genotipo (2df)	0/13/43	1/08/47	-	-	-	-	-
			Test de Cochran armitage	13/99	10/102	0.4439	1	0.5053	-	-
			Alélico Dominante	13/99	10/102	0.4361	1	0.509	1.339	0.5614-3.196
			Recesivo	13/43	09/47.	-	-	-	-	-
rs13069742	A/G	A	Test de Genotipo (2df)	0/13/43	1/08/47	-	-	-	-	-
			Test de Cochran armitage	13/99	10/102	0.4439	1	0.5053	-	-
			Alélico Dominante	13/99	10/102	0.4361	1	0.509	1.339	0.5614-3.196
			Recesivo	13/43	09/47.	-	-	-	-	-
rs35110864	G/A	G	Test de Genotipo (2df)	0/9/47	0/5/51	-	-	-	-	-
			Test de Cochran armitage	9/103	5/107	1.306	1	0.2531	-	-
			Alélico Dominante	9/103	5/107	1.219	1	0.2695	1.87	0.6063-5.767
			Recesivo	09/47.	05/51.	-	-	-	-	-
rs13085367	T/C	T	Test de Genotipo (2df)	1/07/48	0/4/52	-	-	-	-	-
			Test de Cochran armitage	9/103	4/108	1.853	1	0.1734	-	-
			Alélico Dominante	9/103	4/108	2.042	1	0.1531	2.359	0.7047-7.898
			Recesivo	08/48.	04/52.	-	-	-	-	-
rs4443214	T/C	T	Test de Genotipo (2df)	1/08/47	0/8/48	-	-	-	-	-
			Test de Cochran armitage	10/102	8/104	0.2338	1	0.6287	-	-
			Alélico Dominante	10/102	8/104	0.2416	1	0.623	1.275	0.4836-3.359
			Recesivo	09/47.	08/48.	-	-	-	-	-
rs35775079	C/T	C	Test de Genotipo (2df)	1/06/49	0/1/55	-	-	-	-	-
			Test de Cochran armitage	8/104	1/111	4.768	1	0.02899*	-	-
			Alélico Dominante	8/104	1/111	5.672	1	0.01723*	8.538	1.05-69.45
			Recesivo	07/49.	01/55.	-	-	-	-	-
rs11919389	T/C	T	Test de Genotipo (2df)	1/17/35	2/20/34	-	-	-	-	-
			Test de Cochran armitage	19/87	24/88	0.455	1	0.5	-	-
			Alélico Dominante	19/87	24/88	0.4223	1	0.5158	0.8008	0.4094-1.566
			Recesivo	18/35	22/34	-	-	-	-	-
rs343320	G/A	G	Test de Genotipo (2df)	0/4/52	1/08/46.	-	-	-	-	-
			Test de Cochran armitage	4/108	10/100	2.637	1	0.1044	-	-
			Alélico Dominante	4/108	10/100	2.861	1	0.09073	0.3704	0.1126-1.219
			Recesivo	04/52.	09/46.	-	-	-	-	-
rs56162149	C/T	C	Test de Genotipo (2df)	4/16/36	1/17/38	-	-	-	-	-
			Test de Cochran armitage	24/88	19/93	0.6851	1	0.4078	-	-
			Alélico Dominante	24/88	19/93	0.7195	1	0.3963	1.335	0.684-2.605
			Recesivo	20/36	18/38	-	-	-	-	-

Tabla 8: Análisis de asociación genético para COVID-19 severo

SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico en casos	Genotipo alélico en controles	X2	df	Valor p	OR	IC 95%
rs9271609	T/C	T	Test de Genotipo (2df)	9/22/25	2/34/20	-	-	-	-	-
			Test de Cochran armitage	40/72	38/74	0.08757	1	0.7673	-	-
			Alélico Dominante	40/72	38/74	0.07868	1	0.7791	1.082	0.6242-1.875
			Alélico Recesivo	31/25	36/20	-	-	-	-	-
			Alélico Recesivo	09/47.	02/54.	-	-	-	-	-
rs2496644	A/C	C	Geotyping (2df) test Cochran-Armitage trend test	4/22/30	3/23/30	-	-	-	-	-
			Test de Cochran armitage	30/82	29/83	0.02386	1	0.8773	-	-
			Alélico Dominante	30/82	29/83	0.02301	1	0.8794	1.047	0.5777-1.898
			Alélico Recesivo	26/30	26/30	-	-	-	-	-
			Alélico Recesivo	04/52.	03/53.	-	-	-	-	-
rs1886814	A/C	A	Test de Genotipo (2df)	3/20/33	3/20/33	-	-	-	-	-
			Test de Cochran armitage	26/86	26/86	0	1	1	-	-
			Alélico Dominante	26/86	26/86	0	1	1	1	0.5378-1.86
			Alélico Recesivo	23/33	23/33	-	-	-	-	-
			Alélico Recesivo	03/53.	03/53.	-	-	-	-	-
rs28368148	C/G	C	Test de Genotipo (2df)	0/1/55	0/1/55	-	-	-	-	-
			Test de Cochran armitage	1/111	1/111	0	1	1	-	-
			Alélico Dominante	1/111	1/111	0	1	1	1	0.06177-16.19
			Alélico Recesivo	01/55.	01/55.	-	-	-	-	-
			Alélico Recesivo	0/56	0/56	-	-	-	-	-
rs505922	T/C	T	Test de Genotipo (2df)	2/16/38	1/16/39	-	-	-	-	-
			Test de Cochran armitage	20/92	18/94	0.1286	1	0.7199	-	-
			Alélico Dominante	20/92	18/94	0.1268	1	0.7218	1.135	0.5645-2.283
			Alélico Recesivo	18/38	17/39	-	-	-	-	-
			Alélico Recesivo	02/54.	01/55.	-	-	-	-	-
rs529565	T/C	T	Test de Genotipo (2df)	3/19/34	1/16/39	-	-	-	-	-
			Test de Cochran armitage	25/87	18/94	1.421	1	0.2333	-	-
			Alélico Dominante	25/87	18/94	1.41	1	0.235	1.501	0.7661-2.94
			Alélico Recesivo	22/34	17/39	-	-	-	-	-
			Alélico Recesivo	03/53.	01/55.	-	-	-	-	-
rs61882275	G/A	G	Test de Genotipo (2df)	13/25/18	6/32/18	3.439	2	0.1792	-	-
			Test de Cochran armitage	51/61	44/68	0.9348	1	0.3336	-	-
			Alélico Dominante	51/61	44/68	0.8956	1	0.344	1.292	0.7597-2.198
			Alélico Recesivo	38/18	38/18	0	1	1	-	-
			Alélico Recesivo	13/43	06/50.	3.106	1	0.07801	-	-
rs10774671	G/A	A	Test de Genotipo (2df)	4/12/40	3/21/32	-	-	-	-	-
			Test de Cochran armitage	20/92	27/85	1.187	1	0.2759	-	-
			Alélico Dominante	20/92	27/85	1.319	1	0.2507	0.6844	0.3576-1.31
			Alélico Recesivo	16/40	24/32	-	-	-	-	-
			Alélico Recesivo	04/52.	03/53.	-	-	-	-	-
rs2660	G/A	A	Test de Genotipo (2df)	4/09/43	1/18/37	-	-	-	-	-
			Test de Cochran armitage	17/95	20/92	0.2588	1	0.611	-	-
			Alélico Dominante	17/95	20/92	0.2914	1	0.5893	0.8232	0.4058-1.67
			Alélico Recesivo	13/43	19/37	-	-	-	-	-
			Alélico Recesivo	04/52.	01/55.	-	-	-	-	-
rs10850097	C/T	T	Test de Genotipo (2df)	4/11/41	2/20/34	-	-	-	-	-
			Test de Cochran armitage	19/93	24/88	0.6495	1	0.4203	-	-
			Alélico Dominante	19/93	24/88	0.7195	1	0.3963	0.7491	0.3838-1.462
			Alélico Recesivo	15/41	22/34	-	-	-	-	-
			Alélico Recesivo	04/52.	04/54.	-	-	-	-	-

Tabla 8: Análisis de asociación genético para COVID-19 severo

SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico en casos	Genotipo alélico en controles	X2	df	Valor p	OR	IC 95%
rs6489867	C/T	T	Test de Genotipo (2df)	5/10/41	2/22/32	-	-	-	-	-
			Test de Cochran armitage	20/92	26/86	0.8758	1	0.3494	-	-
			Alélico Dominante	20/92	26/86	0.9849	1	0.321	0.7191	0.3743-1.381
			Alélico Recesivo	15/41	24/32	-	-	-	-	-
			Alélico Recesivo	05/51.	04/54.	-	-	-	-	-
rs7955267	C/T	T	Test de Genotipo (2df)	6/10/40	1/22/33	-	-	-	-	-
			Test de Cochran armitage	22/90	24/88	0.09731	1	0.7551	-	-
			Alélico Dominante	22/90	24/88	0.1094	1	0.7408	0.8963	0.4684-1.715
			Alélico Recesivo	16/40	23/33	-	-	-	-	-
			Alélico Recesivo	06/50.	01/55.	-	-	-	-	-
rs56106917	C/D	C	Test de Genotipo (2df)	5/23/28	5/19/32	0.6476	2	0.7234	-	-
			Test de Cochran armitage	33/79	29/83	0.3356	1	0.5624	-	-
			Alélico Dominante	33/79	29/83	0.3568	1	0.5503	1.196	0.6652-2.149
			Alélico Recesivo	28/28	24/32	0.5744	1	0.4485	-	-
			Alélico Recesivo	05/51.	05/51.	0	1	1	-	-
rs9577175	C/T	C	Test de Genotipo (2df)	3/25/28	1/22/33	-	-	-	-	-
			Test de Cochran armitage	31/81	24/88	1.361	1	0.2433	-	-
			Alélico Dominante	31/81	24/88	1.181	1	0.2772	1.403	0.7606-2.589
			Alélico Recesivo	28/28	23/33	-	-	-	-	-
			Alélico Recesivo	03/53.	01/55.	-	-	-	-	-
rs4424872	T/A	A	Test de Genotipo (2df)	1/01/54	2/01/53	-	-	-	-	-
			Test de Cochran armitage	3/109	5/107	0.2979	1	0.5852	-	-
			Alélico Dominante	3/109	5/107	0.5185	1	0.4715	0.589	0.1373-2.526
			Alélico Recesivo	04/54.	03/53.	-	-	-	-	-
			Alélico Recesivo	01/55.	02/54.	-	-	-	-	-
rs117169628	G/A	G	Test de Genotipo (2df)	3/09/44	0/13/43	-	-	-	-	-
			Test de Cochran armitage	15/97	13/99	0.1481	1	0.7003	-	-
			Alélico Dominante	15/97	13/99	0.1633	1	0.6862	1.178	0.5325-2.604
			Alélico Recesivo	12/44.	13/43	-	-	-	-	-
			Alélico Recesivo	03/53.	0/56	-	-	-	-	-
rs79600142	T/C	T	Test de Genotipo (2df)	1/05/50	0/15/41	-	-	-	-	-
			Test de Cochran armitage	7/105	15/97	3.252	1	0.07132	-	-
			Alélico Dominante	7/105	15/97	3.226	1	0.07248	0.4311	0.1686-1.102
			Alélico Recesivo	06/50.	15/41	-	-	-	-	-
			Alélico Recesivo	01/55.	0/56	-	-	-	-	-
rs62054835	A/C	A	Test de Genotipo (2df)	1/04/51	0/15/41	-	-	-	-	-
			Test de Cochran armitage	6/106	15/97	4.249	1	0.03927*	-	-
			Alélico Dominante	6/106	15/97	4.256	1	0.03911*	0.366	0.1366-0.9812
			Alélico Recesivo	05/51.	15/41	-	-	-	-	-
			Alélico Recesivo	01/55.	0/56	-	-	-	-	-
rs112572874	A/G	A	Test de Genotipo (2df)	1/05/50	0/16/40	-	-	-	-	-
			Test de Cochran armitage	7/105	16/96	3.995	1	0.04564*	-	-
			Alélico Dominante	7/105	16/96	3.925	1	0.04758*	0.4	0.1578-1.014
			Alélico Recesivo	06/50.	16/40	-	-	-	-	-
			Alélico Recesivo	01/55.	0/56	-	-	-	-	-
rs1819040	T/A	T	Test de Genotipo (2df)	1/05/50	2/13/41	-	-	-	-	-
			Test de Cochran armitage	7/105	17/95	4.023	1	0.04488*	-	-
			Alélico Dominante	7/105	17/95	4.667	1	0.03075*	0.3725	0.148-0.9376
			Alélico Recesivo	06/50.	15/41	-	-	-	-	-
			Alélico Recesivo	01/5.	04/54.	-	-	-	-	-

Tabla 8: Análisis de asociación genético para COVID-19 severo

SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico en casos	Genotipo alélico en controles	X2	df	Valor p	OR	IC 95%	
rs2532300	T/C	T	Test de Genotipo (2df)	1/05/50	0/15/41	-	-	-	-	-	
			Test de Cochran armitage	7/105	15/97	3.252	1	0.07132	-	-	0.1686-
			Alélico Dominante	7/105	15/97	3.226	1	0.07248	0.4311	-	1.102
			Alélico Recesivo	06/50. 01/55.	15/41 0/56	-	-	-	-	-	-
rs3848456	C/A	C	Test de Genotipo (2df)	2/22/32	1/26/29	-	-	-	-	-	
			Test de Cochran armitage	26/86	28/84	0.1178	1	0.7315	-	-	-
			Alélico Dominante	26/86	28/84	0.0976	1	0.7547	0.907	-	0.4915- 1.674
			Alélico Recesivo	24/32 02/54.	27/29 01/55.	-	-	-	-	-	-
rs77534576	C/T	C	Test de Genotipo (2df)	2/13/41	0/19/37	-	-	-	-	-	
			Test de Cochran armitage	17/95	19/93	0.1407	1	0.7076	-	-	-
			Alélico Dominante	17/95	19/93	0.1324	1	0.716	0.8759	-	0.4289- 1.789
			Alélico Recesivo	15/41 02/54.	19/37 0/56	-	-	-	-	-	-
rs12610495	A/G	A	Test de Genotipo (2df)	7/18/31	3/19/34	-	-	-	-	-	
			Test de Cochran armitage	32/80	25/87	1.021	1	0.3123	-	-	-
			Alélico Dominante	32/80	25/87	1.153	1	0.2829	1.392	-	0.7602- 2.549
			Alélico Recesivo	25/31 07/49.	22/34 03/53.	-	-	-	-	-	-
rs2109069	G/A	G	Test de Genotipo (2df)	7/20/29	3/22/31	-	-	-	-	-	
			Test de Cochran armitage	34/78	28/84	0.7551	1	0.3849	-	-	-
			Alélico Dominante	34/78	28/84	0.8029	1	0.3702	1.308	-	0.7267- 2.353
			Alélico Recesivo	27/29 07/49.	25/31 03/53.	-	-	-	-	-	-
rs2277732	C/A	C	Test de Genotipo (2df)	7/19/30	3/20/33	-	-	-	-	-	
			Test de Cochran armitage	33/79	26/86	1.023	1	0.3119	-	-	-
			Alélico Dominante	33/79	26/86	1.127	1	0.2883	1.382	-	0.7599- 2.512
			Alélico Recesivo	26/30 07/49.	23/33 03/53.	-	-	-	-	-	-
rs4804803	A/G	A	Test de Genotipo (2df)	3/11/42	2/16/38	-	-	-	-	-	
			Test de Cochran armitage	17/95	20/92	0.2588	1	0.611	-	-	-
			Alélico Dominante	17/95	20/92	0.2914	1	0.5893	0.8232	-	0.4058- 1.67
			Alélico Recesivo	14/42 03/53.	18/38 02/54.	-	-	-	-	-	-
rs73510898	G/A	G	Test de Genotipo (2df)	2/01/53	0/5/51	-	-	-	-	-	
			Test de Cochran armitage	5/107	5/107	0	1	1	-	-	-
			Alélico Dominante	5/107	5/107	0	1	1	1	-	0.2813- 3.555
			Alélico Recesivo	03/53. 02/54.	may-51 0/56	-	-	-	-	-	-
rs74956615	T/A	T	Test de Genotipo (2df)	0/2/54	0/7/49	-	-	-	-	-	
			Test de Cochran armitage	2/110	7/105	3.02	1	0.08222	-	-	-
			Alélico Dominante	2/110	7/105	2.894	1	0.08891	0.2727	-	0.05539- 1.343
			Alélico Recesivo	02/54. 0/56	07/49. 0/56	-	-	-	-	-	-
rs34536443	G/C	G	Test de Genotipo (2df)	0/2/54	0/4/52	-	-	-	-	-	
			Test de Cochran armitage	2/110	4/108	0.7044	1	0.4013	-	-	-
			Alélico Dominante	2/110	4/108	0.685	1	0.4079	0.4909	-	0.08808- 2.736
			Alélico Recesivo	02/54. 0/56	04/52. 0/56	-	-	-	-	-	-

Tabla 8: Análisis de asociación genético para COVID-19 severo

SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico en casos	Genotipo alélico en controles	X2	df	Valor p	OR	IC 95%
rs429358	T/C	T	Test de Genotipo (2df)	07/49	2/10/44	-	-	-	-	-
			Test de Cochran armitage	7/105	14/98	2.326	1	0.1272	-	-
			Alélico	7/105	14/98	2.575	1	0.1086	0.4667	-
			Dominante	07/49.	12/44.	-	-	-	-	-
			Recesivo	0/56	02/54.	-	-	-	-	-
rs368565	C/T	T	Test de Genotipo (2df)	19/16/20	9/20/27	5.05	2	0.08007	-	-
			Test de Cochran armitage	54/56	38/74	3.948	1	0.04694 *	-	-
			Alélico	54/56	38/74	5.257	1	0.02185 *	1.878	1.093-3.226
			Dominante	35/20	29/27	1.596	1	0.2064	-	-
			Recesivo	19/36	09/47.	5.021	1	0.02505 *	-	-
rs4801778	G/T	G	Test de Genotipo (2df)	1/12/43	2/13/41	-	-	-	-	-
			Test de Cochran armitage	14/98	17/95	0.3167	1	0.5736	-	-
			Alélico	14/98	17/95	0.337	1	0.5616	0.7983	0.3728-1.71
			Dominante	13/43	15/41	-	-	-	-	-
			Recesivo	01/55.	02/54.	-	-	-	-	-
rs17860115	C/A	C	Test de Genotipo (2df)	16/26/14	12/30/14	0.8571	2	0.6514	-	-
			Test de Cochran armitage	58/54	54/58	0.2857	1	0.593	-	-
			Alélico	58/54	54/58	0.2857	1	0.593	1.154	0.683-1.949
			Dominante	42/14	42/14	0	1	1	-	-
			Recesivo	16/40	12/44.	0.7619	1	0.3827	-	-
rs2300370	G/A	G	Test de Genotipo (2df)	17/26/13	11/30/15	1.714	2	0.4244	-	-
			Test de Cochran armitage	60/52	52/60	1.143	1	0.285	-	-
			Alélico	60/52	52/60	1.143	1	0.285	1.331	0.7874-2.251
			Dominante	43/13	41/15	0.1905	1	0.6625	-	-
			Recesivo	17/39	11/45.	1.714	1	0.1904	-	-
rs2252639	A/G	A	Test de Genotipo (2df)	13/27/16	11/32/13	0.9007	2	0.6374	-	-
			Test de Cochran armitage	53/59	54/58	0.01895	1	0.8905	-	-
			Alélico	53/59	54/58	0.01789	1	0.8936	0.9648	0.5711-1.63
			Dominante	40/16	43/13	0.4188	1	0.5175	-	-
			Recesivo	13/43	11/45.	0.2121	1	0.6451	-	-
rs2236757	A/G	G	Test de Genotipo (2df)	15/24/17	10/32/14	2.433	2	0.2962	-	-
			Test de Cochran armitage	54/58	52/60	0.07184	1	0.7887	-	-
			Alélico	54/58	52/60	0.07163	1	0.789	1.074	0.6357-1.815
			Dominante	39/17	42/14	0.4014	1	0.5263	-	-
			Recesivo	15/41	10/46.	1.287	1	0.2565	-	-
rs2300371	C/T	C	Test de Genotipo (2df)	13/24/19	10/32/14	2.292	2	0.3179	-	-
			Test de Cochran armitage	50/62	52/60	0.07259	1	0.7876	-	-
			Alélico	50/62	52/60	0.072	1	0.7884	0.9305	0.5499-1.575
			Dominante	37/19	42/14	1.074	1	0.3	-	-
			Recesivo	13/43	10/46.	0.4924	1	0.4828	-	-
rs8178521	C/T	C	Test de Genotipo (2df)	7/15/34	2/22/32	-	-	-	-	-
			Test de Cochran armitage	29/83	26/86	0.1957	1	0.6582	-	-
			Alélico	29/83	26/86	0.2169	1	0.6414	1.156	0.6284-2.125
			Dominante	22/34	24/32	-	-	-	-	-
			Recesivo	07/49.	02/54.	-	-	-	-	-
rs147418963	D/AC	D	Test de Genotipo (2df)	2/19/35	0/19/37	-	-	-	-	-
			Test de Cochran armitage	23/89	19/93	0.5289	1	0.4671	-	-
			Alélico	23/89	19/93	0.4689	1	0.4935	1.265	0.645-2.481
			Dominante	21/35	19/37	-	-	-	-	-
			Recesivo	02/54.	0/56	-	-	-	-	-

Tabla 8: Análisis de asociación genético para COVID-19 severo

SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico en casos	Genotipo alélico en controles	X2	df	Valor p	OR	IC 95%
rs2298660	C/T	C	Test de Genotipo (2df)	3/15/38	2/19/35	-	-	-	-	-
			Test de Cochran armitage	21/91	23/89	0.1089	1	0.7413	-	-
			Alélico Dominante	21/91	23/89	0.1131	1	0.7366	0.893	0.4616-1.727
			Alélico Recesivo	18/38	21/35	-	-	-	-	-
			Alélico Recesivo	03/53.	02/54.	-	-	-	-	-
rs2298661	C/A	C	Test de Genotipo (2df)	3/14/39	3/19/34	-	-	-	-	-
			Test de Cochran armitage	20/92	25/87	0.6423	1	0.4229	-	-
			Alélico Dominante	20/92	25/87	0.6952	1	0.4044	0.7565	0.3922-1.459
			Alélico Recesivo	17/39	22/34	-	-	-	-	-
			Alélico Recesivo	03/53.	03/53.	-	-	-	-	-
rs3787946	G/C	C	Test de Genotipo (2df)	3/16/37	3/19/34	-	-	-	-	-
			Test de Cochran armitage	22/90	25/87	0.2291	1	0.6322	-	-
			Alélico Dominante	22/90	25/87	0.2423	1	0.6225	0.8507	0.4466-1.62
			Alélico Recesivo	19/37	22/34	-	-	-	-	-
			Alélico Recesivo	03/53.	03/53.	-	-	-	-	-

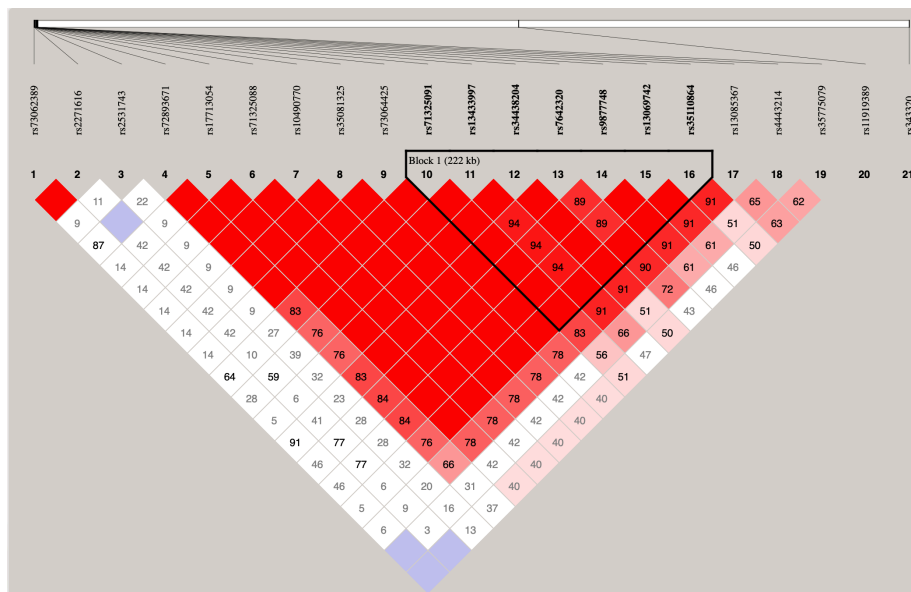
SNP (polimorfismo de nucleótido simple); A2 (Alelo menor); Ref (Alelo de referencia) Alt (Alelo alterno)

*Genotipo con asociación significativa; $p < 0.05$; OR (Odds Ratio); IC (intervalos de confianza del 95%)

10.5.3. Desequilibrio de ligamiento

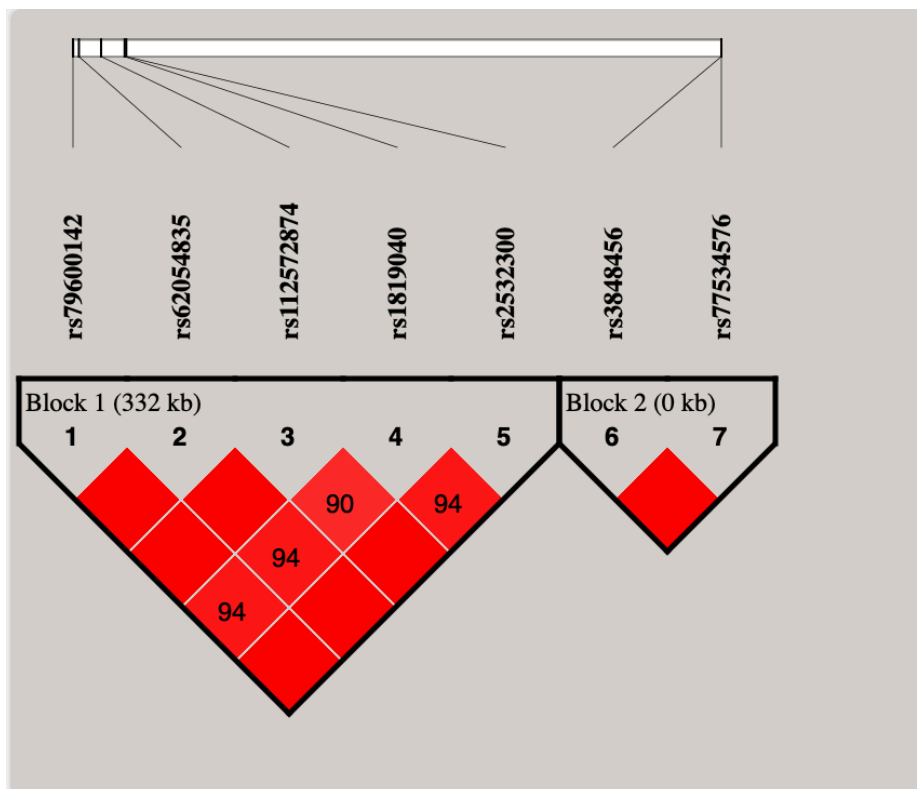
De las variantes con hallazgos significativos se encontró que muchas de ellas estaban localizadas en regiones genómicas cercanas en un mismo cromosoma, por esta razón se evaluó si presentaban desequilibrio de ligamiento. En el cromosoma 3 se identificaron 5 variantes en desequilibrio de ligamiento (rs17713054, rs71325088, rs10490770, rs35081325, rs73064425). (Figura 2) De manera similar en el cromosoma 17 se encontraron 3 variantes en desequilibrio de ligamiento (rs62054835, rs112572874, rs1819040) (Figura 3). Estas observaciones se tuvieron en cuenta para los análisis posteriores pues debido a su cercanía física estas variantes no se segregan de forma independiente, por lo que solamente se incluyó uno de los SNP del bloque de desequilibrio de ligamiento dentro de los modelos predictivos para evitar sobrerepresentación.

Figura 2. Análisis de desequilibrio de ligamiento SNPs Chr 3



Desequilibrio de ligamiento entre los SNPs ubicados en el cromosoma 3 analizados por el software Haploview 4.2. Los cuadrados muestran en su interior el valor de D' . Para los cuadrado rojo intenso en los cuales no se incluyen valores numericos el $D' = 1$

Figura 3. Análisis de desequilibrio de ligamiento SNP Chr 17



Desequilibrio de ligamiento entre los SNPs ubicados en el cromosoma 17 analizados por el software Haploview 4.2

Los cuadrados muestran en su interior el valor de D'. Para los cuadrado rojo intenso en los cuales no se incluyen valores numericos el D' = 1.

10.5.4 Análisis bivariado con presencia de COVID-19 prolongado

Respecto al análisis bivariado entre los polimorfismos genéticos y la presencia de COVID-19 prolongado, se encontraron 4 variantes asociadas significativamente ubicadas en diferentes cromosomas y ninguna de ellas se presentó en desequilibrio de ligamiento. Destacando la variante rs8178521 en el gen *IL10RB* la cual presentó la mayor fuerza de asociación ($p < 0.01$; OR=2.51; 95% CI=1.27-4.94) (Tabla 9).

Tabla 9. Análisis bivariado entre los polimorfismos genéticos con la presencia de COVID-19 prolongado

SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico con COVID-19 prolongado	Genotipo alélico sin COVID-19 prolongado	X2	df	Valor p	OR	IC 95%	
rs114301457	C/T	C	Test de Genotipo (2df)	0/0/66	0/1/45	-	-	-	-	-	
			Test de Cochran								
			armitage	0/132	01/91.	1.448	1	0.2289	-	-	
			Alélico	0/132	01/91.	1.441	1	0.2299	0	-	
			Dominante	0/66	01/45.	-	-	-	-	-	
rs7528026	G/A	G	Test de Genotipo (2df)	0/66	0/46	-	-	-	-	-	
			Test de Cochran	0/4/62	0/1/45	-	-	-	-	-	
			armitage	4/128	01/91.	0.9601	1	0.3272	-	-	
			Alélico	4/128	01/91.	0.9382	1	0.3327	2.844	0.3127-25.86	
			Dominante	04/62.	01/45.	-	-	-	-	-	
rs41264915	A/G	A	Test de Genotipo (2df)	0/66	0/46	-	-	-	-	-	
			Test de Cochran	0/7/59	0/4/42	-	-	-	-	-	
			armitage	7/125	04/88.	0.1117	1	0.7382	-	-	
			Alélico	7/125	04/88.	0.1059	1	0.7448	1.232	0.35-4.336	
			Dominante	07/59.	04/42.	-	-	-	-	-	
rs1123573	A/G	A	Test de Genotipo (2df)	0/66	0/46	-	-	-	-	-	
			Test de Cochran	1/30/35	2/14/30	-	-	-	-	-	
			armitage	32/100	18/74	0.7888	1	0.3745	-	-	
			Alélico	32/100	18/74	0.684	1	0.4082	1.316	0.6861-2.522	
			Dominante	31/35	16/30	-	-	-	-	-	
rs2232354	T/G	T	Test de Genotipo (2df)	01/65.	02/44.	-	-	-	-	-	
			Test de Cochran	6/17/43	4/10/32	-	-	-	-	-	
			armitage	29/103	18/74	0.1485	1	0.7	-	-	
			Alélico	29/103	18/74	0.1891	1	0.6637	1.157	0.5985-2.239	
			Dominante	23/43	14/32	-	-	-	-	-	
rs147509469	G/A	A	Test de Genotipo (2df)	06/60.	04/42.	-	-	-	-	-	
			Test de Cochran	0/0/66	1/01/44	-	-	-	-	-	
			armitage	0/132	03/89.	2.625	1	0.1052	-	-	
			Alélico	0/132	03/89.	4.363	1	0.03673*	0	-	
			Dominante	0/66	02/44.	-	-	-	-	-	
rs73062389	A/G	G	Test de Genotipo (2df)	0/66	01/45.	-	-	-	-	-	
			Test de Cochran	0/6/60	1/04/41	-	-	-	-	-	
			armitage	6/126	06/86.	0.3731	1	0.5413	-	-	
			Alélico	6/126	06/86.	0.4176	1	0.5181	0.6825	0.213-2.187	
			Dominante	06/60.	05/41.	-	-	-	-	-	
rs2271616	G/T	T	Test de Genotipo (2df)	0/66	01/45.	-	-	-	-	-	
			Test de Cochran	2/12/52	1/09/36	-	-	-	-	-	
			armitage	16/116	11/81.	0.001243	1	0.9719	-	-	
			Alélico	16/116	11/81.	0.001387	1	0.9703	1.016	0.448-2.303	
			Dominante	14/52	10/36.	-	-	-	-	-	
rs2531743	G/A	A	Test de Genotipo (2df)	02/84	01/45.	-	-	-	-	-	
			Test de Cochran	0/31/35	4/20/22	-	-	-	-	-	
			armitage	31/101	28/64	1.633	1	0.2013	-	-	
			Alélico	31/101	28/64	1.35	1	0.2453	0.7016	0.3853-1.277	
			Dominante	31/35	24/22	-	-	-	-	-	
rs72893671	T/A	T	Test de Genotipo (2df)	0/66	04/42.	-	-	-	-	-	
			Test de Cochran	0/10/56	0/5/41	-	-	-	-	-	
			armitage	10/122	05/87.	0.4285	1	0.5127	-	-	
			Alélico	10/122	05/87.	0.3977	1	0.5283	1.426	0.4709-4.32	
			Dominante	10/56.	05/41.	-	-	-	-	-	

Tabla 9: Análisis de asociación genético para COVID-19 prolongado

SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico con COVID-19 prolongado	Genotipo alélico sin COVID-19 prolongado	X2	df	Valor p	OR	IC 95%
rs17713054	G/A	G	Test de Genotipo (2df)	0/6/60	0/4/42	-	-	-	-	-
			Test de Cochran armitage	6/126	04/88	0.005208	1	0.9425	-	-
			Alélico	6/126	04/88	0.004965	1	0.9438	1.048	0.2872-3.821
			Dominante	06/60	04/42	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs71325088	T/C	T	Test de Genotipo (2df)	0/6/60	0/4/42	-	-	-	-	-
			Test de Cochran armitage	6/126	04/88	0.005208	1	0.9425	-	-
			Alélico	6/126	04/88	0.004965	1	0.9438	1.048	0.2872-3.821
			Dominante	06/60	04/42	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs10490770	T/C	T	Test de Genotipo (2df)	0/6/60	0/4/42	-	-	-	-	-
			Test de Cochran armitage	6/126	04/88	0.005208	1	0.9425	-	-
			Alélico	6/126	04/88	0.004965	1	0.9438	1.048	0.2872-3.821
			Dominante	06/60	04/42	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs11385942	D/A	D	Test de Genotipo (2df)	0/7/59	0/4/42	-	-	-	-	-
			Test de Cochran armitage	7/125	04/88	0.1117	1	0.7382	-	-
			Alélico	7/125	04/88	0.1059	1	0.7448	1.232	0.35-4.336
			Dominante	07/59	04/82	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs35081325	A/T	A	Test de Genotipo (2df)	0/6/60	0/4/42	-	-	-	-	-
			Test de Cochran armitage	6/126	04/88	0.005208	1	0.9425	-	-
			Alélico	6/126	04/88	0.004965	1	0.9438	1.048	0.2872-3.821
			Dominante	06/60	04/42	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs73064425	C/T	C	Test de Genotipo (2df)	0/6/60	0/4/42	-	-	-	-	-
			Test de Cochran armitage	6/126	04/88	0.005208	1	0.9425	-	-
			Alélico	6/126	04/88	0.004965	1	0.9438	1.048	0.2872-3.821
			Dominante	06/60	04/42	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs71325091	G/A	G	Test de Genotipo (2df)	0/9/57	0/4/42	-	-	-	-	-
			Test de Cochran armitage	9/123	04/88	0.6449	1	0.4219	-	-
			Alélico	9/123	04/88	0.6052	1	0.4366	1.61	0.4804-5.394
			Dominante	09/57	04/42	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs13433997	T/C	T	Test de Genotipo (2df)	0/15/51	0/6/39	-	-	-	-	-
			Test de Cochran armitage	15/117	06/84	1.539	1	0.2147	-	-
			Alélico	15/117	06/84	1.378	1	0.2404	1.795	0.6687-4.818
			Dominante	15/51	06/39	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs34438204	T/C	T	Test de Genotipo (2df)	0/9/57	0/5/41	-	-	-	-	-
			Test de Cochran armitage	9/123	05/87	0.1897	1	0.6631	-	-
			Alélico	9/123	05/87	0.1771	1	0.6739	1.273	0.4124-3.93
			Dominante	09/57	05/41	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs7642320	A/G	A	Test de Genotipo (2df)	0/19/47	0/10/36	-	-	-	-	-
			Test de Cochran armitage	19/113	10/82	0.7019	1	0.4022	-	-
			Alélico	19/113	10/82	0.5975	1	0.4395	1.379	0.6092-3.12
			Dominante	19/47	10/36	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs9877748	A/G	A	Test de Genotipo (2df)	0/14/52	1/07/38	-	-	-	-	-
			Test de Cochran armitage	14/118	09/83	0.04061	1	0.8403	-	-
			Alélico	14/118	09/83	0.0399	1	0.8417	1.094	0.4524-2.646
			Dominante	14/52	08/38	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs13069742	A/G	A	Test de Genotipo (2df)	0/14/52	1/07/38	-	-	-	-	-
			Test de Cochran armitage	14/118	09/83	0.04061	1	0.8403	-	-
			Alélico	14/118	09/83	0.0399	1	0.8417	1.094	0.4524-2.646
			Dominante	14/52	08/38	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs35110864	G/A	G	Test de Genotipo (2df)	0/9/57	0/5/41	-	-	-	-	-
			Test de Cochran armitage	9/123	05/87	0.1897	1	0.6631	-	-
			Alélico	9/123	05/87	0.1771	1	0.6739	1.273	0.4124-3.93
			Dominante	09/57	05/41	-	-	-	-	-
			Recesivo	0/66	0/46	-	-	-	-	-
rs13085367	T/C	T	Test de Genotipo (2df)	1/07/58	0/4/42	-	-	-	-	-
			Test de Cochran armitage	9/123	04/88	0.5493	1	0.4586	-	-
			Alélico	9/123	04/88	0.6052	1	0.4366	1.61	0.4804-5.394
			Dominante	08/58	04/42	-	-	-	-	-
			Recesivo	0/65	0/46	-	-	-	-	-
rs4443214	T/C	T	Test de Genotipo (2df)	1/12/53	0/4/42	-	-	-	-	-
			Test de Cochran armitage	14/118	04/88	2.78	1	0.09543	-	-
			Alélico	14/118	04/88	2.873	1	0.09006	2.61	0.8306-8.202
			Dominante	13/53	04/42	-	-	-	-	-
			Recesivo	0/65	0/46	-	-	-	-	-
rs35775079	C/T	C	Test de Genotipo (2df)	1/06/59	0/1/45	-	-	-	-	-
			Test de Cochran armitage	8/124	01/91	2.923	1	0.08732	-	-
			Alélico	8/124	01/91	3.478	1	0.06221	5.871	0.7215-47.77
			Dominante	07/59	01/45	-	-	-	-	-
			Recesivo	0/65	0/46	-	-	-	-	-
rs11919389	T/C	T	Test de Genotipo (2df)	1/19/44	2/18/25	-	-	-	-	-
			Test de Cochran armitage	21/107	22/68	2.323	1	0.1274	-	-
			Alélico	21/107	22/68	2.156	1	0.142	0.6066	0.3102-1.186
			Dominante	20/44	20/25	-	-	-	-	-
			Recesivo	0/63	02/43	-	-	-	-	-

Tabla 9: Análisis de asociación genético para COVID-19 prolongado										
SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico con COVID-19 prolongado	Genotipo alélico sin COVID-19 prolongado	X2	df	Valor p	OR	IC 95%
rs343320	G/A	G	Test de Genotipo (2df)	1/06/59	0/6/39	-	-	-	-	-
			Test de Cochran	8/124	06/84.	0.03066	1	0.861	-	-
			armitage	8/124	06/84.	0.03327	1	0.8553	0.9032	0.3024-2.698
			Alélico	07/59.	06/39.	-	-	-	-	-
			Dominante	01/65.	0/45	-	-	-	-	-
rs56162149	C/T	C	Test de Genotipo (2df)	4/18/44	1/15/30	-	-	-	-	-
			Test de Cochran	26/106	17/75	0.04943	1	0.8241	-	-
			armitage	26/106	17/75	0.05191	1	0.8198	1.082	0.5487-2.134
			Alélico	22/44	16/30	-	-	-	-	-
			Dominante	04/62.	01/45.	-	-	-	-	-
rs9271609	T/C	T	Test de Genotipo (2df)	9/31/26	2/25/19	-	-	-	-	-
			Test de Cochran	49/83	29/63	0.8336	1	0.3612	-	-
			armitage	49/83	29/63	0.749	1	0.3668	1.283	0.7296-2.254
			Alélico	40/26	27/19	-	-	-	-	-
			Dominante	09/57.	04/44.	-	-	-	-	-
rs2496644	A/C	C	Test de Genotipo (2df)	4/30/32	3/15/28	-	-	-	-	-
			Test de Cochran	38/94	21/71	1.03	1	0.3102	-	-
			armitage	38/94	21/71	0.9932	1	0.319	1.367	0.7385-2.53
			Alélico	34/32	18/28	-	-	-	-	-
			Dominante	04/62.	03/43.	-	-	-	-	-
rs1886814	A/C	A	Test de Genotipo (2df)	3/29/34	3/11/32	-	-	-	-	-
			Test de Cochran	35/97	17/75	1.968	1	0.1607	-	-
			armitage	35/97	17/75	1.965	1	0.161	1.592	0.8285-3.059
			Alélico	32/34	14/32	-	-	-	-	-
			Dominante	03/63.	03/43.	-	-	-	-	-
rs28368148	C/G	C	Test de Genotipo (2df)	0/1/65	0/1/45	-	-	-	-	-
			Test de Cochran	1/131	0/1/91.	0.06707	1	0.7956	-	-
			armitage	1/131	0/1/91.	0.06647	1	0.7965	0.6947	0.04289-11.25
			Alélico	01/65.	01/45.	-	-	-	-	-
			Dominante	0/66	0/46	-	-	-	-	-
rs505922	T/C	T	Test de Genotipo (2df)	1/20/45	2/12/32	-	-	-	-	-
			Test de Cochran	22/110	16/76	0.0205	1	0.8862	-	-
			armitage	22/110	16/76	0.02021	1	0.887	0.95	0.4684-1.927
			Alélico	21/45	14/32	-	-	-	-	-
			Dominante	01/65.	02/44.	-	-	-	-	-
rs529565	T/C	T	Test de Genotipo (2df)	2/23/41	2/12/32	-	-	-	-	-
			Test de Cochran	27/105	16/76	0.3304	1	0.5654	-	-
			armitage	27/105	16/76	0.328	1	0.5669	1.221	0.6156-2.424
			Alélico	25/41	14/32	-	-	-	-	-
			Dominante	02/64.	02/44.	-	-	-	-	-
rs61882275	G/A	G	Test de Genotipo (2df)	13/33/20	6/24/16	0.9018	2	0.6371	-	-
			Test de Cochran	59/73	36/56	0.7179	1	0.3968	-	-
			armitage	59/73	36/56	0.6878	1	0.4069	1.257	0.7317-2.16
			Alélico	46/20	30/16	0.2494	1	0.6175	-	-
			Dominante	13/53	06/40.	0.8519	1	0.356	-	-
rs10774671	G/A	A	Test de Genotipo (2df)	4/16/46	3/17/26	-	-	-	-	-
			Test de Cochran	24/108	23/69	1.368	1	0.2422	-	-
			armitage	24/108	23/69	1.52	1	0.2176	0.6667	0.3492-1.273
			Alélico	20/46	20/26	-	-	-	-	-
			Dominante	04/62.	03/43.	-	-	-	-	-
rs2660	G/A	A	Test de Genotipo (2df)	3/12/51	2/15/29	-	-	-	-	-
			Test de Cochran	18/114	19/73	1.719	1	0.1898	-	-
			armitage	18/114	19/73	1.935	1	0.1642	0.6066	0.2987-1.232
			Alélico	15/51	17/29	-	-	-	-	-
			Dominante	03/63.	02/44.	-	-	-	-	-
rs10850097	C/T	T	Test de Genotipo (2df)	4/14/48	2/17/27	-	-	-	-	-
			Test de Cochran	22/110	21/71	1.197	1	0.2739	-	-
			armitage	22/110	21/71	1.326	1	0.2495	0.6762	0.3466-1.319
			Alélico	18/48	19/27	-	-	-	-	-
			Dominante	04/62.	02/44.	-	-	-	-	-
rs6489867	C/T	T	Test de Genotipo (2df)	5/14/47	2/18/26	-	-	-	-	-
			Test de Cochran	24/108	22/70	0.9704	1	0.3246	-	-
			armitage	24/108	22/70	1.091	1	0.2962	0.7071	0.3684-1.357
			Alélico	19/47	20/26	-	-	-	-	-
			Dominante	05/61.	02/44.	-	-	-	-	-
rs7955267	C/T	T	Test de Genotipo (2df)	5/15/46	2/17/27	-	-	-	-	-
			Test de Cochran	25/107	21/71	0.4463	1	0.5041	-	-
			armitage	25/107	21/71	0.5019	1	0.4787	0.7899	0.4111-1.518
			Alélico	20/46	19/27	-	-	-	-	-
			Dominante	05/61.	02/44.	-	-	-	-	-
rs56106917	C/D	C	Test de Genotipo (2df)	7/23/36	3/19/24	-	-	-	-	-
			Test de Cochran	37/95	25/67	0.01868	1	0.8913	-	-
			armitage	37/95	25/67	0.01986	1	0.8879	1.044	0.5751-1.894
			Alélico	30/36	22/24	-	-	-	-	-
			Dominante	07/59.	03/43.	-	-	-	-	-
rs9577175	C/T	C	Test de Genotipo (2df)	4/31/31	0/16/30	-	-	-	-	-
			Test de Cochran	39/93	16/76	4.984	1	0.02558	-	-
			armitage	39/93	16/76	4.323	1	0.0376*	1.992	1.034-3.839
			Alélico	35/31	16/30	-	-	-	-	-
			Dominante	04/62.	0/46	-	-	-	-	-
rs4424872	T/A	A	Test de Genotipo (2df)	1/01/64	2/01/43	-	-	-	-	-
			Test de Cochran	3/129	05/87.	0.9042	1	0.3417	-	-
			armitage	3/129	05/87.	1.574	1	0.2096	0.4047	0.09426-1.737
			Alélico	02/64.	03/43.	-	-	-	-	-
			Dominante	01/65.	02/44.	-	-	-	-	-

Tabla 9. Análisis de asociación genético para COVID-19 prolongado

SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico con COVID-19 prolongado	Genotipo alélico sin COVID-19 prolongado	X2	df	Valor p	OR	IC 95%
rs117169628	G/A	G	Test de Genotipo (2df)	2/14/50	1/08/37	-	-	-	-	-
			Test de Cochran armitage	18/114	10/82.	0.3443	1	0.5574	-	-
			Alélico	18/114	10/82.	0.3794	1	0.5379	1.295	0.5683-2.95
			Dominante	16/50	09/37.	-	-	-	-	-
rs79600142	T/C	T	Test de Genotipo (2df)	1/08/57	0/12/34	-	-	-	-	-
			Test de Cochran armitage	10/122	12/80.	1.845	1	0.1744	-	-
			Alélico	10/122	12/80.	1.83	1	0.1761	0.5464	0.2255-1.324
			Dominante	09/57.	12/34.	-	-	-	-	-
rs62054835	A/C	A	Test de Genotipo (2df)	1/08/57	0/11/35	-	-	-	-	-
			Test de Cochran armitage	10/122	11/81.	1.223	1	0.2689	-	-
			Alélico	10/122	11/81.	1.225	1	0.2685	0.6036	0.2451-1.486
			Dominante	09/57.	11/35.	-	-	-	-	-
rs112572874	A/G	A	Test de Genotipo (2df)	1/09/56	0/12/34	-	-	-	-	-
			Test de Cochran armitage	11/121	12/80.	1.329	1	0.249	-	-
			Alélico	11/121	12/80.	1.305	1	0.2532	0.6061	0.2551-1.44
			Dominante	10/56.	12/34.	-	-	-	-	-
rs1819040	T/A	T	Test de Genotipo (2df)	1/08/57	2/10/34	-	-	-	-	-
			Test de Cochran armitage	10/122	14/78	2.853	1	0.09121	-	-
			Alélico	10/122	14/78	3.309	1	0.06889	0.4567	0.1933-1.079
			Dominante	09/57.	12/34.	-	-	-	-	-
rs2532300	T/C	T	Test de Genotipo (2df)	1/08/57	0/12/34	-	-	-	-	-
			Test de Cochran armitage	10/122	12/80.	1.845	1	0.1744	-	-
			Alélico	10/122	12/80.	1.83	1	0.1761	0.5464	0.2255-1.324
			Dominante	09/57.	12/34.	-	-	-	-	-
rs3848456	C/A	C	Test de Genotipo (2df)	3/27/36	0/21/25	-	-	-	-	-
			Test de Cochran armitage	33/99	21/71	0.169	1	0.681	-	-
			Alélico	33/99	21/71	0.14	1	0.7082	1.127	0.6025-2.108
			Dominante	30/36	21/25	-	-	-	-	-
rs77534576	C/T	C	Test de Genotipo (2df)	2/18/46	0/14/32	-	-	-	-	-
			Test de Cochran armitage	22/110	14/78	0.08972	1	0.7645	-	-
			Alélico	22/110	14/78	0.08442	1	0.7714	1.114	0.5369-2.313
			Dominante	20/46	14/32	-	-	-	-	-
rs12610495	A/G	A	Test de Genotipo (2df)	6/24/36	4/13/29	-	-	-	-	-
			Test de Cochran armitage	36/96	21/71	0.5003	1	0.4794	-	-
			Alélico	36/96	21/71	0.565	1	0.4522	1.268	0.6824-2.356
			Dominante	30/36	17/29	-	-	-	-	-
rs2109069	G/A	G	Test de Genotipo (2df)	7/27/32	3/15/28	-	-	-	-	-
			Test de Cochran armitage	41/91	21/71	1.727	1	0.1888	-	-
			Alélico	41/91	21/71	1.836	1	0.1754	1.523	0.8271-2.805
			Dominante	34/32	18/28	-	-	-	-	-
rs2277732	C/A	C	Test de Genotipo (2df)	6/27/33	4/12/30	-	-	-	-	-
			Test de Cochran armitage	39/93	20/72	1.544	1	0.214	-	-
			Alélico	39/93	20/72	1.703	1	0.1919	1.51	0.8116-2.808
			Dominante	33/33	16/30	-	-	-	-	-
rs4804803	A/G	A	Test de Genotipo (2df)	4/15/47	1/12/33	-	-	-	-	-
			Test de Cochran armitage	23/109	14/78	0.1701	1	0.6801	-	-
			Alélico	23/109	14/78	0.1915	1	0.6617	1.176	0.5693-2.428
			Dominante	19/47	13/33	-	-	-	-	-
rs73510898	G/A	G	Test de Genotipo (2df)	1/04/61	1/02/43	-	-	-	-	-
			Test de Cochran armitage	6/126	04/88.	0.003619	1	0.952	-	-
			Alélico	6/126	04/88.	0.004965	1	0.9438	1.048	0.2872-3.821
			Dominante	05/61.	03/43.	-	-	-	-	-
rs74956615	T/A	T	Test de Genotipo (2df)	0/4/62	0/5/41	-	-	-	-	-
			Test de Cochran armitage	4/128	05/87.	0.8483	1	0.357	-	-
			Alélico	4/128	05/87.	0.8128	1	0.3673	0.5438	0.142-2.082
			Dominante	04/62.	05/41.	-	-	-	-	-
rs34536443	G/C	G	Test de Genotipo (2df)	0/6	0/2/44	-	-	-	-	-
			Test de Cochran armitage	4/128	02/90.	0.1568	1	0.6921	-	-
			Alélico	4/128	02/90.	0.1525	1	0.6961	1.406	0.2521-7.843
			Dominante	04/62.	02/44.	-	-	-	-	-
rs429358	T/C	T	Test de Genotipo (2df)	2/10/54	0/7/39	-	-	-	-	-
			Test de Cochran armitage	14/118	07/85.	0.518	1	0.4717	-	-
			Alélico	14/118	07/85.	0.5733	1	0.449	1.441	0.5576-3.722
			Dominante	12/54.	07/39.	-	-	-	-	-
rs429358	T/C	T	Test de Genotipo (2df)	0/2/64.	0/46	-	-	-	-	-
			Test de Cochran armitage	0/2/64.	0/46	-	-	-	-	-
			Alélico	0/2/64.	0/46	-	-	-	-	-
			Dominante	0/2/64.	0/46	-	-	-	-	-
rs429358	T/C	T	Test de Genotipo (2df)	20/21/24	8/15/23	3	2	0.2232	-	-

Tabla 9: Análisis de asociación genético para COVID-19 prolongado											
SNP	Alelo Ref/Alt	Alelo menor (A2)	Modelo	Genotipo alélico con COVID-19 prolongado	Genotipo alélico sin COVID-19 prolongado	X2	df	Valor p	OR	IC 95%	
rs368565	C/T	T	Test de Genotipo (2df)	20/21/24	8/15/23	3	2	0.2232	-	-	
			Test de Cochran								
			armitage	61/69	31/61	2.917	1	0.08767	-	-	
			Alélico	61/69	31/61	3.884	1	0.04874**	1.74	1.001-3.023	
			Dominante	41/24	23/23	1.887	1	0.1696	-	-	
rs4801778	G/T	G	Test de Genotipo (2df)	20/45	08/38.	2.556	1	0.1099	-	-	
			Test de Cochran	1/14/51	2/11/33	-	-	-	-	-	
			armitage	16/116	15/77	0.7477	1	0.3872	-	-	
			Alélico	16/116	15/77	0.7956	1	0.3724	0.708	0.3308-1.516	
			Dominante	15/51	13/33	-	-	-	-	-	
rs17860115	C/A	C	Test de Genotipo (2df)	15/37/14	13/19/14	2.435	2	0.296	-	-	
			Test de Cochran								
			armitage	67/65	45/47	0.07378	1	0.7859	-	-	
			Alélico	67/65	45/47	0.07378	1	0.7859	1.077	0.6321-1.834	
			Dominante	52/14	32/14	1.23	1	0.2675	-	-	
rs2300370	G/A	G	Test de Genotipo (2df)	15/51	13/33	0.4427	1	0.5058	-	-	
			Test de Cochran	15/37/14	13/19/14	2.435	2	0.296	-	-	
			armitage	67/65	45/47	0.07378	1	0.7859	-	-	
			Alélico	67/65	45/47	0.07378	1	0.7859	1.077	0.6321-1.834	
			Dominante	52/14	32/14	1.23	1	0.2675	-	-	
rs2252639	A/G	A	Test de Genotipo (2df)	13/38/15	11/21/14	1.578	2	0.4542	-	-	
			Test de Cochran								
			armitage	64/68	43/49	0.07012	1	0.7912	-	-	
			Alélico	64/68	43/49	0.06622	1	0.7969	1.073	0.6293-1.828	
			Dominante	51/15	32/14	0.8392	1	0.3596	-	-	
rs2236757	A/G	G	Test de Genotipo (2df)	13/53	11/35.	0.2862	1	0.5927	-	-	
			Test de Cochran	14/34/18	11/22/13	0.1719	2	0.9176	-	-	
			armitage	62/70	44/48	0.016	1	0.8994	-	-	
			Alélico	62/70	44/48	0.01595	1	0.8995	0.9662	0.567-1.647	
			Dominante	48/18	33/13	0.01322	1	0.9085	-	-	
rs2300371	C/T	C	Test de Genotipo (2df)	14/52	11/35.	0.114	1	0.7356	-	-	
			Test de Cochran	14/35/17	9/21/16	1.08	2	0.5827	v	-	
			armitage	63/69	39/53	0.6275	1	0.4283	-	-	
			Alélico	63/69	39/53	0.6224	1	0.4302	1.241	0.7257-2.121	
			Dominante	49/17	30/16	1.062	1	0.3027	-	-	
rs8178521	C/T	C	Test de Genotipo (2df)	14/52	09/37.	0.04505	1	0.8319	-	-	
			Test de Cochran	8/25/33	1/12/33	-	-	-	-	-	
			armitage	41/91	14/78	6.628	1	0.01004*	-	-	
			Alélico	41/91	14/78	7.346	1	0.006721	2.51	1.274-4.944	
			Dominante	33/33	13/33	-	-	-	-	-	
rs147418963	D/AC	D	Test de Genotipo (2df)	08/58.	01/45.	-	-	-	-	-	
			Test de Cochran	2/20/44	0/18/28	-	-	-	-	-	
			armitage	24/108	18/74	0.07683	1	0.7816	-	-	
			Alélico	24/108	18/74	0.06811	1	0.7941	0.9136	0.4633-1.801	
			Dominante	22/44	18/28	-	-	-	-	-	
rs2298660	C/T	C	Test de Genotipo (2df)	02/64.	0/46	-	-	-	-	-	
			Test de Cochran	3/19/44	2/15/29	-	-	-	-	-	
			armitage	25/107	19/73	0.09704	1	0.7554	-	-	
			Alélico	25/107	19/73	0.1008	1	0.7509	0.8977	0.4609-1.741	
			Dominante	22/44	17/29	-	-	-	-	-	
rs2298661	C/A	C	Test de Genotipo (2df)	03/63.	02/44.	-	-	-	-	-	
			Test de Cochran	4/18/44	2/15/29	-	-	-	-	-	
			armitage	26/106	19/73	0.02847	1	0.866	-	-	
			Alélico	26/106	19/73	0.03081	1	0.8607	0.9424	0.4859-1.821	
			Dominante	22/44	17/29	-	-	-	-	-	
rs3787946	G/C	C	Test de Genotipo (2df)	02/62.	02/44.	-	-	-	-	-	
			Test de Cochran	4/20/42	2/15/29	-	-	-	-	-	
			armitage	28/104	19/73	0.009694	1	0.9216	-	-	
			Alélico	28/104	19/73	0.01025	1	0.9193	1.034	0.5374-1.991	
			Dominante	24/42	17/29	-	-	-	-	-	

SNP (polimorfismo de nucleótido simple); A2 (Alelo menor); Ref (Alelo de referencia)
Alt (Alelo alterno)

*Genotipo con asociación significativa; $p < 0.05$; OR (Odds Ratio); IC (intervalos de confianza del 95%)

10.6 Análisis de variantes moleculares en genes candidato

El panel de construido para la caracterización genética de la población estudiada incluyó 66 genes relacionados biológicamente con la vía molecular de infección o con la respuesta del huésped ante la misma. Todas las variantes identificadas con sus respectivos análisis genético-poblacionales se encuentran disponibles en la Tabla suplementaria 4. El análisis de este panel de genes permitió la identificación de un total de 288 variantes genéticas de las cuales el 17.4% (n=50) corresponde a variantes de pérdida de función, por su sigla en inglés LoF (Loss-of-function), incluyendo en esta categoría mutaciones tipo frameshift, sin sentido o de sitios canónicos de splicing. El 82.6% restante (n=238) corresponden a mutaciones tipo missense en donde se cambia un aminoácido. De estas el 8.4% (n=20) son mutaciones potencialmente patógenas con puntaje REVEL >0.5 y el 0.4% (n=1) mutación probablemente patogénica usando el método de clasificación de la ACMG. De los genes analizados, 4 (*APOE*, *ICAM5*, *PLSCR1* y *TAC4*) de ellos presentaron exclusivamente variantes tipo LoF el resto de los genes presentan predominantemente mutaciones tipo missense.

Respecto a la proporción de variantes presentes entre casos y controles, se evidenciaron un total de 284 variantes en los casos y 296 en los controles. Las variantes LoF se distribuyeron entre 43 pacientes clasificados como casos vs 37 pacientes clasificados como controles, ($p=0.42$). El promedio de variantes de este tipo por paciente en el grupo de los casos fue mayor que el grupo de los controles, con 0.75 y 0.65 variantes LoF por individuo para el grupo de casos y controles, respectivamente. Con relación a las variantes genéticas missense clasificadas como probablemente patogénicas por el predictor *in silico* REVEL (>0.5) estas se presentaron en 22 casos y 33 controles ($p=0.16$). La distribución de estas variantes por genes y pacientes se encuentran presentadas en la tabla 10. Respecto a la identificación de variantes patogénicas o probablemente patogénicas usando el método de clasificación de la ACMG usando Franklin genoox solo se encontró una variante probablemente patogénica, la cual no ha sido previamente reportada en la literatura (c.1429C>T) en el gen *PDE4A*. Se identificaron un total de 97 variantes nuevas de las cuales el 43.3% (n=42) corresponden a variantes tipo LoF y 6.2% (n=6) a variantes con potencial patogénico según el predictor *in silico*, estas se muestran en la tabla suplementaria 4.

Tabla 10. Comparación de variantes LoF y missense patogénicas entre casos y controles por gen. Esta tabla presenta el número de casos o controles con variantes tipo LoF o Missense patogénicas (REVEL puntaje >0.5) por cada gen.

Tabla 10. Comparación de variantes LoF y missense patogénicas entre casos y controles por gen						
Gen	Casos LoF	Casos Missense patogénicas	Casos Total	Controles LoF	Controles Missense patogénicas	Controles Total
<i>APOE</i>	1	0	1	0	0	0
<i>APOL1</i>	0	0	0	0	0	0
<i>ARHGAP27</i>	0	0	0	0	0	0
<i>ARL17B</i>	2	0	2	1	0	1
<i>ATP11A</i>	1	1	2	1	0	1
<i>BCL11A</i>	0	0	0	0	0	0
<i>CCR3</i>	0	0	0	0	0	0
<i>CCR5</i>	0	0	0	0	0	0
<i>CCR9</i>	0	0	0	0	0	0
<i>CD14</i>	0	0	0	0	1	1
<i>CENPS</i>	0	0	0	0	0	0
<i>CFAP73</i>	0	0	0	0	0	0
<i>DPP4</i>	2	1	3	3	0	3
<i>DPP9</i>	0	0	0	0	0	0
<i>FCGR2A</i>	1	0	1	2	0	2
<i>FDX2</i>	0	0	0	0	0	0
<i>FOXP4</i>	1	2	3	0	2	2
<i>FURIN</i>	0	0	0	0	0	0
<i>FUT2</i>	1	13	14	0	17	17
<i>FYCO1</i>	0	0	0	0	0	0
<i>HSD17B14</i>	1	0	1	0	1	1
<i>ICAM1</i>	0	0	0	0	0	0
<i>ICAM3</i>	0	0	0	0	0	0
<i>ICAM5</i>	0	0	0	1	0	1
<i>IFITM3</i>	0	0	0	0	0	0

Tabla 10. Comparación de variantes LoF y missense patogénicas entre casos y controles por gen

Gen	Casos LoF	Casos Missense patogénicas	Casos Total	Controles LoF	Controles Missense patogénicas	Controles Total
<i>IFNA10</i>	0	0	0	0	0	0
<i>IFNAR1</i>	0	0	0	0	0	0
<i>IFNAR2</i>	0	0	0	3	0	3
<i>IRF3</i>	0	0	0	0	0	0
<i>IRF7</i>	0	0	0	0	0	0
<i>KANSL1</i>	0	0	0	0	0	0
<i>KLRC2</i>	0	0	0	0	0	0
<i>LZTFL1</i>	1	0	1	0	0	0
<i>MAPT</i>	0	0	0	1	0	1
<i>MUC1</i>	5	0	5	6	0	6
<i>MX1</i>	1	0	1	0	0	0
<i>NAPSA</i>	0	0	0	0	1	1
<i>NCOA4</i>	0	0	0	0	0	0
<i>NUCB1</i>	0	0	0	0	0	0
<i>OAS1</i>	0	0	0	0	0	0
<i>OAS3</i>	2	0	2	0	2	2
<i>OCLN</i>	0	0	0	2	0	2
<i>PDE4A</i>	0	0	0	0	1	1
<i>PIGN</i>	1	0	1	0	1	1
<i>PLSCR1</i>	1	0	1	1	0	1
<i>PPP1R15A</i>	0	0	0	0	0	0
<i>RPL24</i>	0	1	1	0	0	0
<i>SLC30A5</i>	1	1	2	0	0	0
<i>SLC6A20</i>	0	0	0	0	0	0
<i>SPDEF</i>	0	0	0	0	0	0
<i>TAC4</i>	5	0	5	4	0	4
<i>TBK1</i>	2	0	2	0	0	0

Tabla 10. Comparación de variantes LoF y missense patogénicas entre casos y controles por gen

Gen	Casos LoF	Casos Missense patogénicas	Casos Total	Controles LoF	Controles Missense patogénicas	Controles Total
<i>THBS3</i>	0	1	1	0	0	0
<i>TICAM1</i>	0	0	0	1	0	1
<i>TLR3</i>	9	0	9	0	1	1
<i>TLR7</i>	0	0	0	0	0	0
<i>TMEM65</i>	1	0	1	0	0	0
<i>TMPRSS2</i>	0	0	0	0	0	0
<i>TRIM46</i>	0	0	0	0	0	0
<i>TYK2</i>	0	2	2	3	5	8

10.7 Modelos predictivos de COVID-19

Las variables clínicas y genéticas identificadas con asociación significativa fueron incorporadas para la construcción de los modelo de regresión logística binaria. Para cada uno de los desenlaces, severidad de COVID-19 y presencia de COVID-19 prolongado, se desarrollaron 3 modelos predictivos con distintos parámetros (clínico, genético y mixto). Estos corresponden a:

1. Modelo clínico el cual incorporó solo variables clínicas no genéticas.
2. Modelo genético en el cual se involucraron solo los SNP
3. Modelo mixto el cual integraron variables tanto clínicas como genéticas

Se utilizó el método *Stepwise* tipo *backward* para seleccionar el mejor modelo utilizando el valor de AIC. De los tres modelos propuestos, tanto para predicción de severidad de COVID-19 como para predicción de COVID-19 prolongado, los mejores modelos predictivos según el valor de AIC y poder de discriminación fueron los modelos mixtos, el cual para el caso de severidad obtuvo un AIC:118.4 (AUC 0.86; 95% CI: 0.7844-0.9271) y para COVID-19 prolongado un AIC: 120.798 (AUC 0.83 95% CI: 0.74-0.91). La comparación de los tres modelos respecto a sus variables, valores p y OR se encuentra en las Tablas 11 y 12, y Figura 4.

Tabla 11. Comparación de los diferentes modelos predictivos Clínico, Genético y Mixto para severidad de COVID-19.

Tabla 11: Comparación de los diferentes modelos predictivos Clínico, Genético y Mixto para severidad de COVID-19.															
Predictores	Modelo Clínico					Modelo Genético					Modelo Mixto				
	OR	SE	IC	Estadístico	Valor p	OR	SE	IC	Estadístico	Valor p	OR	SE	IC	Estadístico	Valor p
Intercepto	0.02	0.03	0.00-0.47	-2.40	0.017	7.74	9.29	0.74-81.39	1.70	0.088	0.13	0.30	0.00-11.83	-0.89	0.375
Sexo masculino	2.62	1.16	1.10-6.24	2.18	0.029						2.72	1.36	1.02-7.27	1.99	0.046
IMC	1.16	0.07	1.03-1.29	2.55	0.011						1.17	0.08	1.02-1.33	2.29	0.022
ComorbNinguna	0.49	0.23	0.20-123	-1.52	0.128						0.37	0.20	0.13-1.05	-1.87	0.062
Snp5						0.35	0.17	0.14-0.89	-2.22	0.042	0.42	0.22	0.15-1.18	-1.64	0.102
Snp49						4.71	2.97	1.37-16.21	0.014	5.87	5.87	4.11	1.49-23.15	2.53	0.011
Snp14						0.06	0.06	0.01-0.52	0.011	0.03	0.03	0.04	0.00-0.37	-2.78	0.006
Observaciones	107	107	107												
Desviación	126.443	126.826	104.43												
AIC	134.443	134.826	118.437												
Log Razon de verosimilitud	-63.221	-63.413	-52.219												

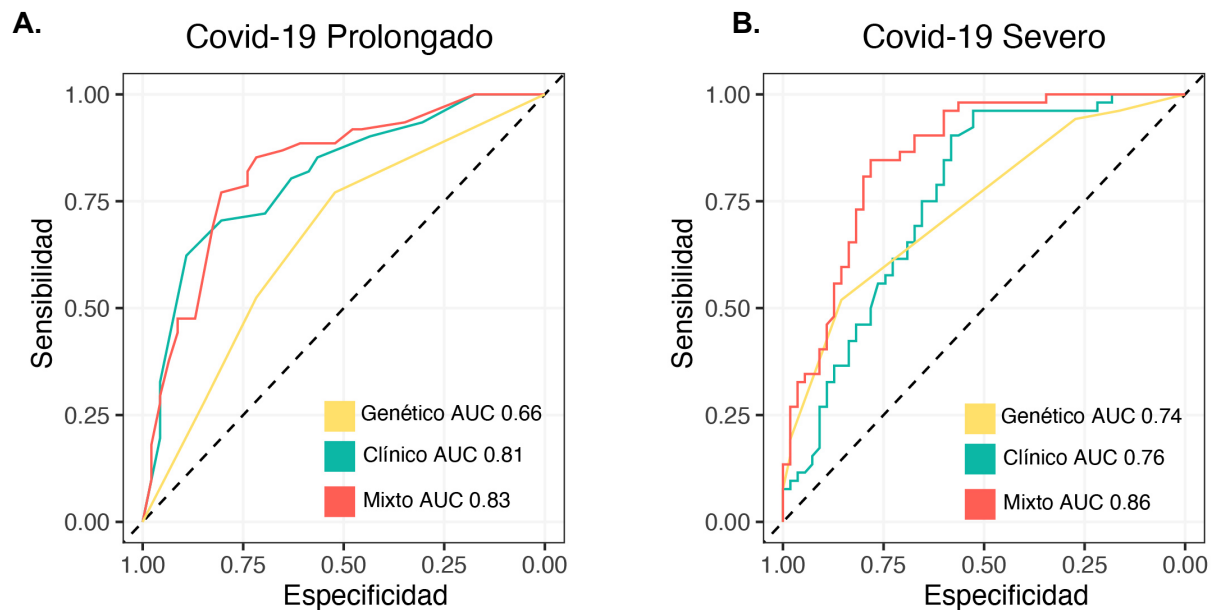
OR (Odds Ratio); SE (Error estándar); IC (Intervalo de confianza del 95%); AIC (Criterio de Akaike), IMC (Índice de masa corporal), SNP5 (rs2232354), SNP49 (rs1819040), SNP14 (rs11385942)

Tabla 12. Comparación de los diferentes modelos predictivos Clínico, Genético y Mixto para presencia de COVID-19 prolongado.

Tabla 12: Comparación de los diferentes modelos predictivos Clínico, Genético y Mixto para presencia de COVID-19 prolongado															
Predictores	Modelo Clínico					Modelo Genético					Modelo Mixto				
	OR	SE	IC	Estadístico	Valor p	OR	SE	IC	Estadístico	Valor p	OR	SE	IC	Estadístico	Valor p
Intercepto	0.08	0.06	0.02-0.31	-3.68	<0.001	3.36	1.34	1.54-7.34	3.04	0.002	0.18	0.14	0.04-0.79	-2.27	0.023
Pérdida del gusto	4.12	2.06	1.55-11	2.83	0.005						3.23	1.67	1.17-8.89	2.27	0.023
Fiebre > 38°C	2.04	0.99	0.79-5.29	1.47	0.141						2.25	1.13	0.84-6.02	1.61	0.107
Fatiga/Cansancio	3.82	2.28	1.19-12.28	2.25	0.024						3.21	1.98	0.96-10.75	1.89	0.059
Severidad	4.42	2.24	1.64-11.92	2.94	0.003						4.98	2.63	1.77-14.02	3.04	0.002
Snp43						0.54	0.22	0.24-1.20	-1.51	0.131					
Snp68						0.38	0.16	0.17-0.87	-2.28	0.023	0.35	0.18	0.13-0.95	-2.06	0.040
Observaciones	107	107	107												
Desviación	113.223	137.488	108.798												
AIC	123.223	143.488	120.798												
Log Razon de verosimilitud	-56.612	-68.744	-54.399												

OR (Odds Ratio); SE (Error estándar); IC (Intervalo de confianza del 95%); AIC (Criterio de Akaike), Severidad (Caso ó Control), SNP43 (rs9577175), SNP (rs8178521)

Figura 4. Comparación curvas ROC de los diferentes modelos predictivos para COVID-19 prologando o para severidad de COVID-19.



A) Curvas ROC modelos predictivos Clínico, Genético y Mixto para presencia de COVID-19 prolongado

B) Curvas ROC modelos predictivos Clínico, Genético y Mixto para severidad de COVID-19

10.7.1 Modelo predictivo de severidad

El modelo seleccionado (mixto) incluye variables clínicas tales como el sexo masculino, IMC, “Ninguna Comorbilidad” y los polimorfismos genéticos SNP rs2232354, SNP rs11385942 y SNP rs1819040 pertenecientes a los genes *IL1RN*, *LZTFL1* y *KANSL1*. La ecuación resultante es la siguiente.

Score ajustado

$$= \frac{1}{1 + e^{-(-2.0430 + (0.9999 * Sexo) + 0.1541(IMC) - 1.0025(No\ comorbilidad) - 0.8604 \left(\frac{Snp5WT}{alt}\right) - 3.3642 \left(\frac{Snp14WT}{alt}\right) + 1.7703(Snp49WT/alt)}}$$

Ecuación 1. Modelo de predicción de riesgo de severidad de COVID-19. Donde “Sexo” hace referencia al sexo biológico masculino, “IMC” hace referencia al índice de masa corporal como variable numérica continua, “ComorbNinguna” a la ausencia total de antecedentes clínicos de riesgo en el paciente, el SNP5 (WT/Alt) y SNP14 (WT/Alt) que ilustran la presencia de los alelos de referencia para los polimorfismos rs2232354 y rs11385942 respectivamente, comportándose como protectores. Y, finalmente el SNP49, que corresponde a la presencia del alelo de referencia, el cual se comporta como de riesgo rs1819040.

Se validó la calidad y supuestos del modelo. Identificando ausencia de colinealidad (factor de inflación de la varianza < 1.2). Homocedasticidad (Breusch Pagan p = 0.0673) y calibración (Hosmer Lemeshow p = 0.095). Sin embargo, se encontró autocorrelación de los errores (Durbin-Watson p < 0.01)

Por lo cual, el perfil de un paciente que más probabilidad de presentar el desenlace de severidad es el siguiente: ser de sexo masculino, tener un elevado IMC y tener el polimorfismo de riesgo SNP49 (rs1819040).

10.7.2 Modelo predictivo de COVID-19 prolongado

El modelo seleccionado corresponde al modelo mixto el cual incorpora las variables clínicas pérdida del gusto, fiebre, fatiga y severidad de cuadro clínico inicial de COVID-19 y el SNP rs8178521 en el gen *IL10RB*. La ecuación predictiva resultante es la siguiente

$$\text{Adjusted score} = \frac{1}{1 + e^{-(-1.72 + 1.60(\text{covid severo}) + 1.17(\text{pérdida del gusto}) + 1.16(\text{Fatiga}) + 0.80(\text{Fiebre}) + 1.06(\text{WT/Alt}))}}$$

Donde COVID-19 severo hace referencia a la necesidad de hospitalización en piso o Unidad de Cuidados Intensivos durante la enfermedad; pérdida del gusto, fatiga y fiebre hacen referencia a los síntomas presentados durante el cuadro clínico agudo y "WT/Alt" hace referencia a la presencia del alelo alterno para el polimorfismo rs8178521. Se validaron los supuestos del modelo no colinealidad (inflación de varianza < 1.2), homocedasticidad (Breusch-Pagan p=0.79) e independencia de los errores (Durbin-Watson p=0.15), se determinó adecuada calibración del modelo (Hosmer y Lemeshow p= 0.58).

A manera de ejemplo entonces, el perfil que más probabilidad de presentar COVID-19 prolongado es un paciente que haya tenido: COVID-19 severo, pérdida del gusto entre su sintomatología, fiebre y fatiga y la presencia del alelo alterno para el plimofismo de riesgo (rs8178521).

11. DISCUSIÓN

EL SARS-CoV2 se transformó en un periodo de meses en un evento de salud pública que afectó todos los aspectos de la vida a nivel global, considerandose como una de las peores pandemias del siglo XXI. En la actualidad, los virus respiratorios como el COVID-19 tienen una capacidad de expansión y transmisibilidad superior a otros microorganismos, lo que les confiere una prioridad en el ámbito de la investigación médica, abordando sus aspectos clínicos, virales y del huésped que expliquen la severidad, progresión y pronóstico de la enfermedad.

En este estudio nos enfocamos en caracterizar el impacto de las variables clínicas y del huésped con la severidad de la infección por COVID-19 y el desarrollo de COVID-19 prolongado en una muestra de la población colombiana. A través del presente estudio identificamos múltiples factores de riesgo genéticos y no genéticos, los cuales fueron incorporados a dos modelos predictivos, uno para severidad y otro para COVID-19 prolongado.

A pesar de que factores virales, tales como variantes no fueron consideradas en este proyecto, evidenciamos que durante el periodo en el que se realizó este estudio hubo alta prevalencia de linajes considerados en inglés como VBM (variantes en monitorización), la menor clasificación de riesgo epidemiológico según la SIG (del inglés "SARS CoV-2- interagency group) y el CDC, tanto B.1 como B.1.621, B.1.621.1 (Mu), al igual que B.1.67.2 y AY (Delta) (100) (23) Por lo cual creemos firmemente en que la integridad de los datos clínicos y genéticos no fue comprometida por estas variables virales.

Se han reportado múltiples factores no genéticos asociados con un mal pronóstico de la infección, entre ellos se encuentran el sexo masculino, la edad avanzada y múltiples comorbilidades (101). Dada la naturaleza de la variable edad como factor de riesgo, realizamos una corrección de esta a través de la selección de casos pareados como medida de ajuste. Concordante con la literatura internacional, hemos encontrado asociación significativa con el sexo masculino, la presencia de obesidad, el índice de masa corporal y DMT2. Se ha evidenciado el papel de la diabetes mellitus tipo II como un generador de daño endotelial, un estado proinflamatorio y glucotoxicidad, aumentando así el riesgo de eventos tromboembólicos (102), lo cual explicaría su rol como factor de riesgo. Altos índices de masa corporal se han visto asociados a mayor expresión de correceptores *ACE2* y furina a nivel de tejido adiposo, lo que se asocia a mayores tasas de replicación viral, aumentando el riesgo de severidad por la infección (103). Otras comorbilidades analizadas no mostraron asociación significativa, sin embargo, la presencia conjunta de estas mostró un aumento significativo en el riesgo, concordante con previos reportes (72).

Los síntomas respiratorios (disnea, tos y odinofagia) al igual que los síntomas sistémicos (fiebre y fatiga) han mostrado asociaciones significativas con cuadros severos de COVID-19 (104), lo que se explica por la invasión del virión al tracto respiratorio inferior, lo cual afecta el parénquima pulmonar, un efecto que es mediado por citoquinas, factores proinflamatorios y quimioatrayentes (105). Sin embargo, su interpretación debe ser cuidadosa, ya que los casos fueron construidos con definiciones basadas en la presencia de dificultad respiratoria, cuya principal sintomatología es la disnea. La invasión del parénquima pulmonar puede progresar posteriormente a una diseminación sistémica del virus, generando alteraciones multiorgánicas relacionadas con desenlaces clínicos más severos (105). De manera interesante, nuestro estudio identificó la anosmia como un factor protector, lo cual se ha reportado previamente en otros estudios retrospectivos (106). Sin embargo,

resaltamos que la anosmia es más comúnmente encontrada en variantes como Alfa y Beta, que son consideradas variantes de interés (VOI), y que tuvieron una mayor prevalencia al inicio de la pandemia, por lo cual se debe interpretar con cuidado la influencia de la anosmia como un factor protector. Finalmente, así como fue reportado por Peghin y colaboradores (107) y concordante con nuestros hallazgos, los fenotipos severos de la infección mostraron una mayor prevalencia de secuelas y síntomas sugestivos de COVID-19 prolongado.

En nuestra muestra, ninguna comorbilidad o variable socio-demográfica tuvo asociación significativa con la presencia de COVID-19 prolongado, hallazgo no concordante con reportes previos, en los cuales la presencia de múltiples comorbilidades han tenido un impacto significativo (64). Esto es explicable dado el amplio espectro de estudios realizados en muestras con medias de edades mucho más altas que la nuestra (108) (64) (109). Sin embargo, varios síntomas sistémicos y respiratorios durante la infección se encontraron asociados a COVID-19 prolongado, un fenómeno reportado previamente en la literatura e incluyen disnea, tos, fiebre, ageusia, entre otros (65). Se ha propuesto que estos signos y síntomas agudos se pueden prolongar en el tiempo y comprometer otros órganos conduciendo a un proceso inflamatorio sistémico, de daño endotelial, pulmonar y del SNC que provoca finalmente inflamación crónica (64).

Respecto a los factores genéticos, este trabajo corresponde al primer estudio genómico en evaluar la presencia de variantes genéticas de potencial impacto en la severidad de COVID-19 y presencia de COVID-19 prolongado en población colombiana. Se identificaron 50 variantes de pérdida de función distribuidas en diferentes genes implicados en la ruta molecular de infección por SARS-CoV2 y en la respuesta inmunológica del huésped frente a la infección. De manera interesante este tipo de variantes con efecto deletéreo fueron identificadas exclusivamente en los genes *APOE*, *ICAM5*, *PLSCR1* y *TAC4*. Este hallazgo es congruente con varios estudios previos que han destacado el impacto de variantes moleculares en estos genes en el desenlace clínico y la severidad de COVID-19 (110) (111) (112). *APOE* codifica para una proteína con amplia distribución tisular y múltiples funciones, implicadas entre otras en la regulación de la inflamación, lo que suprime la proliferación de los linfocitos T, y regula las funciones de los macrófagos (113). La variante molecular identificada en este gen no ha sido descrita previamente en la literatura (c.-24+204_-24+214delGTGGGGTGGGG), y fue detectada en un caso severo de COVID-19. Dada su potencial implicación funcional es plausible estimar que está relacionada con el desenlace clínico de este paciente. Estudios previos en modelos murinos con SARS-CoV-2 han demostrado el impacto biológico de variantes genéticas patogénicas en este gen. La presencia de la variante rs7412, la cual corresponde al alelo *APOE4* se ha asociado con cargas virales aumentadas y respuesta inmunológica suprimida. Adicionalmente esta variante se relacionó con una menor supervivencia de los pacientes afectados por esta entidad (114).

En general, el patrón de distribución de las variantes de pérdida de función identificadas en este estudio fue mayor en los casos (n=43) que en los controles (n=37). A pesar de que esta tendencia no alcanzó significancia estadística (p=0.42), estudios previos han mostrado la importancia de este tipo de variantes en la respuesta inmunológica del huésped y la susceptibilidad a COVID-19 severo o crítico (115) (116). Estos hallazgos en conjunto reafirman la importancia clínica de las variantes de pérdida de función en la progresión de la infección por SARS-CoV. Las variantes de pérdida de función tienen gran impacto dado incluyen variantes nonsense donde se incluye un codón de parada prematuro, frameshift donde se corre el marco de lectura o de splicing dando lugar a proteínas truncadas o aberrantes no funcionales.

Con relación a las variantes missense probablemente patogénicas, estas se encontraron en menor proporción en los casos (n=22) que en los controles (n=33). Esta diferencia no fue significativa (p= 0.162) y difícil llegar a una conclusión respecto a estos hallazgos dado que al tratarse de predicciones *in silico* se requiere análisis funcionales que corroboren su impacto biológico. Adicionalmente otro factor limitante de este estudio corresponde al tamaño de muestra. Las variantes missense corresponden a cambios moleculares donde se cambia un aminoácido por otro que dado sus características intrínsecas pueden impactar en la estructura de la proteína e impactar en su función.

El análisis de asociación de polimorfismos genéticos y el cuadro severo de la enfermedad, permitió la identificación de 13 variantes asociadas significativamente con el desarrollo de COVID-19 severo en la población estudiada. Notablemente, cuatro de las variantes con mayor fuerza de asociación (rs11385942, rs35081325, rs73064425 y rs10490770) se encuentran ubicadas en el gen *LZTFL1*, para los cuales el alelo menor confiere un riesgo significativo de desarrollar cuadros clínicos severos/críticos. La asociación de este gen con el desenlace clínico, se ha descrito ampliamente en varios estudios de diversos países predominantemente de la región europea (38) (18) (15); posicionándolo como uno de los marcadores genéticos clave en la evolución de la infección y respuesta del huésped. Su implicación biológica radica en su alta expresión en las vías respiratorias superiores donde tiene numerosas funciones en la respuesta inmunitaria del huésped ante agentes infecciosos como el SARS-CoV-2. De manera similar, otro de los polimorfismos más relevantes de este estudio fue rs35775079. Este polimorfismo está ubicado en una región intrónica del gen receptor de quimioquinas *CCR3* implicado en la migración de macrófagos, NK y monocitos. Estudios previos han reportado la importancia clínica de este gen, existiendo una relación una mayor expresión génica y alteraciones de los niveles circulantes de células inmunes en casos severos y críticos comparados con pacientes asintomáticos (117) (118).

Respecto a los polimorfismos asociados a la presencia de COVID-19 prolongado, el polimorfismo de mayor asociación corresponde al rs8178521 en el gen *IL10RB*. Este gen codifica para la subunidad beta del receptor de interleucina-10. Análisis previos han determinado que la sobreexpresión *in vitro* de este gen se asocia con un aumento de la carga viral (115). Adicionalmente, estudios en modelos *in vivo* han demostrado que la sobreexpresión de *IL10RB* aumenta la susceptibilidad de sobreinfecciones pulmonares bacterianas letales a través de un aumento de la señalización de IL-10 posviral que atenúa la respuesta inmune (119). Es plausible pensar que, al estar relacionado con una mayor carga viral y una respuesta inmune atenuada, variantes moleculares en este gen se asocien con un estado de COVID-19 prolongado, secundario a una inadecuada depuración viral y permanencia de remanentes del virus circulante. De manera importante, aunque varios factores clínicos y sociodemográficos se han asociado al desarrollo de esta condición, existe información escasa acerca de factores genéticos, por lo que se requieren más estudios que permitan comprender las bases biológicas de esta entidad y poder optimizar estrategias de prevención o manejo oportuno.

El presente estudio permitió conocer un panorama general del perfil genético de riesgo para la severidad de COVID-19 y COVID-19 prolongado en una muestra de la población Colombiana; contribuyendo el avance y ampliación del conocimiento en esta área. Adicionalmente suministran un punto de partida para estudios posteriores en el campo de la genética y medicina personalizada del país.

Con relación a nuestro último objetivo, el análisis realizado permitió la identificación de factores de riesgo genéticos y no genéticos los cuales fueron incorporados a modelos predictivos de riesgo, construidos a través de regresiones logísticas multivariadas. Nuestro modelo mixto de severidad incluyó variables como el sexo masculino, índice de masa corporal, la presencia de ninguna comorbilidad, y dos polimorfismos genéticos (rs2232354 y rs11385942). Estas variables han sido reportadas como factores de riesgo independientes (18,120). Sin embargo, solo dos de estas variables (sexo masculino y rs11385942) han sido utilizadas en modelos predictivos de riesgo clínico-genéticos que evalúan la severidad de COVID-19 en el contexto latinoamericano (38). La utilidad de modelos de este tipo en el contexto clínico ha sido previamente reportado en otros países (118), destacándose la importancia de factores genéticos para mejorar el poder discriminatorio(29).

Los modelos mixtos presentados en este estudio constituyen herramientas de predicción sólida, esto teniendo en cuenta los valores de AUC obtenidos exclusivamente con factores clínicos o genéticos. Mientras el modelo de COVID prolongado cumple con todos los supuestos de modelamiento regresivo, el modelo de severidad no cumple con el supuesto de independencia de errores. A pesar de utilizar herramientas de ajuste como el procedimiento de Cochrane Orcutt o Hildreth Lu, no fue posible realizar la corrección de este supuesto, por lo que concluimos la presencia de una o más variables que no fueron medidas, y que pueden influenciar

significativamente el desenlace de la enfermedad. Entre estas variables podemos encontrar tiempo y espacio, las cuales, a pesar de ser importantes en enfermedades infecciosas, no fueron inicialmente consideradas.

Las variables de tiempo y espacio son imperativas a tener en cuenta para la construcción de modelamientos estadísticos predictivos de enfermedades infecciosas (119). Hipotetizamos que debido a que la recolección de datos ocurrió durante el segundo y tercer pico infeccioso en Colombia, la temporalidad epidemiológica influyó en el fenómeno de autocorrelación de los datos. De igual manera, no se tuvo en cuenta el concepto de autocorrelación espacial, primordial en el estudio de la epidemiología de enfermedades infecciosas (123) y que explica cómo la recolección de estos datos en la ciudad de Bogotá, como epicentro de la pandemia nacional, pudo influenciar también la aparición de autocorrelación.

12. LIMITACIONES

Este estudio presenta algunas limitaciones que se mencionan a continuación: A) El tamaño de muestra. Aunque estimamos que este tamaño de muestra es adecuado para identificar los SNP con tamaño de efecto mediano o grande, puede ser insuficiente para identificar SNP con tamaños de efecto menores como en el caso de variantes raras frecuencia alélica mínima $MAF < 0.01$. B) Aunque se intentó controlar el sesgo de memoria intrínseco de los pacientes en quienes se realizó la encuesta tiempo posterior a la infección, aun existe una posibilidad de que un factor haya influenciado algunos resultados sobre todo respecto a los síntomas manifestados. C) El modelo de regresión de severidad de COVID-19 debe ser empleado con precaución dado que no se logró ajustar este modelo para el cumplimiento del supuesto de incorrelación de los errores. D) Dado el tamaño de muestra estos modelos no se validarán por lo que son necesarios estudios adicionales que puedan determinar su validez.

Este estudio no tuvo en cuenta en su diseño original la epidemiología molecular del virus y sus variantes durante el tiempo de recolección de la muestra, ni tampoco durante el seguimiento, dado que su objetivo principal estuvo centrado en la identificación de los factores genéticos asociados a la severidad por SARS-CoV-2 y COVID-19 prolongado.

13. CONCLUSIONES

Nuestro estudio analizó la relación entre factores genéticos y no genéticos con el desenlace de severidad de SARS-CoV-2 y COVID-19 prolongado en la población colombiana, demostrando una asociación positiva entre la gravedad de la infección, presencia de COVID prolongado y múltiples factores clínicos y polimorfismos genéticos ubicados en genes responsables de la señalización inmunológica, celular y la respuesta contra microorganismos.

Carterizamos el perfil de riesgo genómico y de susceptibilidad clínica a la severidad de COVID-19 y COVID-19 prolongado en una muestra de la población colombiana. Respondiendo entonces a nuestros objetivos, encontramos múltiples polimorfismos implicados en el desenlace clínico de COVID-19 (rs11385942, rs35775079) previamente reportados en la literatura. Identificamos de igual manera variantes moleculares potencialmente patogénicas en genes candidatos como APOE y TLR3, genes imperativos para el correcto funcionamiento inmunológico y señalización celular.

De igual manera confirmamos la influencia y asociación de comorbilidades clínicas como Diabetes Mellitus Tipo2, aparición de signos y síntomas específicos como disnea o fatiga con el desenlace de severidad clínica por COVID-19, al igual que la alta prevalencia de trastornos neuropsiquiátricos en pacientes con COVID-19 prolongado.

Resaltamos de igual manera la utilidad de estudios genómicos y epidemiológicos para la construcción de herramientas de medicina traslacional dirigidas a COVID-19. La incorporación de dichas herramientas en un contexto clínico podría permitir la identificación de individuos o poblaciones de alto riesgo, así como intervenciones individualizadas dirigidas a reducir la morbi-mortalidad y optimizar el uso de recursos de salud limitados.

Consideramos imperativo y crítico el papel de la temporalidad epidemiológica y la autocorrelación espacial como conceptos que deben tenerse en cuenta a la hora de realizar modelamientos epidemiológicos de enfermedades infecciosas, para así construir mejores modelos predictivos de adecuada validez.

14. BIBLIOGRAFÍA:

1. OMS. Emergencias. 2022 [citado 19 de diciembre de 2022]. Coronavirus disease (COVID-19) pandemic. Disponible en: <https://www.who.int/europe/emergencias/situations/covid-19>
2. COVID - Coronavirus Statistics - Worldometer [Internet]. [citado 2 de mayo de 2023]. Disponible en: <https://www.worldometers.info/coronavirus/>
3. John Hopkins University. CoVID19 dashboard. 2022 [citado 19 de diciembre de 2022]. Coronavirus world Map. Disponible en: <https://gisanddata.maps.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6>
4. Lefkowitz EJ, Dempsey DM, Hendrickson RC, Orton RJ, Siddell SG, Smith DB. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Res.* 4 de enero de 2018;46(D1):D708-17.
5. Yuki K, Fujiogi M, Koutsogiannaki S. COVID-19 pathophysiology: A review. *Clin Immunol.* junio de 2020;215:108427.
6. Hu B, Guo H, Zhou P, Shi ZL. Characteristics of SARS-CoV-2 and COVID-19.

Nat Rev Microbiol. marzo de 2021;19(3):141-54.

7. Adil MT, Rahman R, Whitelaw D, Jain V, Al-Taani O, Rashid F, et al. SARS-CoV-2 and the pandemic of COVID-19. *Postgrad Med J*. febrero de 2021;97(1144):110-6.
8. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19) [Internet]. 2022 [citado 19 de diciembre de 2022]. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). Disponible en: [https://www.who.int/publications-detail-redirect/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-\(covid-19\)](https://www.who.int/publications-detail-redirect/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19))
9. Carfi A, Bernabei R, Landi F, for the Gemelli Against COVID-19 Post-Acute Care Study Group. Persistent Symptoms in Patients After Acute COVID-19. *JAMA*. 11 de agosto de 2020;324(6):603.
10. Logue JK, Franko NM, McCulloch DJ, McDonald D, Magedson A, Wolf CR, et al. Sequelae in Adults at 6 Months After COVID-19 Infection. *JAMA Netw Open*. 19 de febrero de 2021;4(2):e210830.
11. Ovsyannikova IG, Haralambieva IH, Crooke SN, Poland GA, Kennedy RB. The role of host genetics in the immune response to SARS-CoV-2 and COVID-19 susceptibility and severity. *Immunol Rev*. julio de 2020;296(1):205-19.
12. Clouston SAP, Luft BJ, Sun E. Clinical risk factors for mortality in an analysis of 1375 patients admitted for COVID treatment. *Sci Rep*. 3 de diciembre de 2021;11(1):23414.
13. Forster P, Forster L, Renfrew C, Forster M. Phylogenetic network analysis of SARS-CoV-2 genomes. *Proc Natl Acad Sci*. 28 de abril de 2020;117(17):9241-3.
14. Velavan TP, Pallerla SR, Rüter J, Augustin Y, Kreamsner PG, Krishna S, et al. Host genetic factors determining COVID-19 susceptibility and severity. *eBioMedicine*. octubre de 2021;72:103629.
15. The GenOMICC Investigators, The ISARIC4C Investigators, The COVID-19 Human Genetics Initiative, 23andMe Investigators, BRACOVIC Investigators, GenCOVID Investigators, et al. Genetic mechanisms of critical illness in COVID-19. *Nature*. 4 de marzo de 2021;591(7848):92-8.
16. Wang F, Huang S, Gao R, Zhou Y, Lai C, Li Z, et al. Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. *Cell Discov*. 10 de noviembre de 2020;6(1):83.
17. Mousa M, Vurivi H, Kannout H, Uddin M, Alkaabi N, Mahboub B, et al. Genome-wide association study of hospitalized COVID-19 patients in the United Arab Emirates. *eBioMedicine*. diciembre de 2021;74:103695.
18. The Severe Covid-19 GWAS Group. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *N Engl J Med*. 15 de octubre de 2020;383(16):1522-34.
19. COVID-19 Host Genetics Initiative, COVID-19 Host Genetics Initiative Leadership, Niemi MEK, Karjalainen J, Liao RG, Neale BM, et al. Mapping the human genetic architecture of COVID-19. *Nature*. 16 de diciembre de 2021;600(7889):472-7.
20. Proal AD, VanElzakker MB. Long COVID or Post-acute Sequelae of COVID-

- 19 (PASC): An Overview of Biological Factors That May Contribute to Persistent Symptoms. *Front Microbiol.* 23 de junio de 2021;12:698169.
21. Weekly epidemiological update on COVID-19 [Internet]. 2022 [citado 19 de diciembre de 2022]. Weekly epidemiological update on COVID-19 - 30 November 2022. Disponible en: <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---30-november-2022>
22. Reportes de Situación COVID-19: Colombia [Internet]. 2022 [citado 19 de diciembre de 2022]. Reportes de Situación COVID-19: Colombia | OPS/OMS | Organización Panamericana de la Salud. Disponible en: <https://www.paho.org/es/reportes-situacion-covid-19-colombia>
23. COVID19 en Colombia [Internet]. 2022 [citado 19 de diciembre de 2022]. Noticias coronavirus-casos. Disponible en: <https://www.ins.gov.co/Noticias/Paginas/coronavirus-casos.aspx>
24. Booth A, Reed AB, Ponzio S, Yassaee A, Aral M, Plans D, et al. Population risk factors for severe disease and mortality in COVID-19: A global systematic review and meta-analysis. Madeddu G, editor. *PLOS ONE.* 4 de marzo de 2021;16(3):e0247461.
25. Yildirim Z, Sahin OS, Yazar S, Bozok Cetintas V. Genetic and epigenetic factors associated with increased severity of Covid-19. *Cell Biol Int.* junio de 2021;45(6):1158-74.
26. Suh S, Lee S, Gym H, Yoon S, Park S, Cha J, et al. A systematic review on papers that study on Single Nucleotide Polymorphism that affects coronavirus 2019 severity. *BMC Infect Dis.* diciembre de 2022;22(1):47.
27. COVID19 hg [Internet]. 2022 [citado 20 de diciembre de 2022]. COVID-19 Host Genetics Initiative. Disponible en: <https://www.covid19hg.org/partners/>
28. Secolin R, de Araujo TK, Gonsales MC, Rocha CS, Naslavsky M, Marco LD, et al. Genetic variability in COVID-19-related genes in the Brazilian population. *Hum Genome Var.* 2 de abril de 2021;8(1):15.
29. Dite GS, Murphy NM, Allman R. An integrated clinical and genetic model for predicting risk of severe COVID-19: A population-based case-control study. Novelli G, editor. *PLOS ONE.* 16 de febrero de 2021;16(2):e0247205.
30. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* 12 de marzo de 2020;579(7798):270-3.
31. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol.* marzo de 2019;17(3):181-92.
32. Xiao K, Zhai J, Feng Y, Zhou N, Zhang X, Zou JJ, et al. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature.* 9 de julio de 2020;583(7815):286-9.
33. Shi Y, Wang G, Cai X peng, Deng J wen, Zheng L, Zhu H hong, et al. An overview of COVID-19. *J Zhejiang Univ-Sci B.* mayo de 2020;21(5):343-60.
34. Castañeda S, Patiño LH, Muñoz M, Ballesteros N, Guerrero-Araya E, Paredes-Sabja D, et al. Evolution and Epidemic Spread of SARS-CoV-2 in Colombia: A Year into the Pandemic. *Vaccines.* 30 de julio de 2021;9(8):837.

35. Jimenez-Silva C, Rivero R, Douglas J, Bouckaert R, Villabona-Arenas J, Atkins K, et al. Genomic epidemiology of circulating SARS-CoV-2 variants during first two years of the pandemic in Colombia [Internet]. *Genetic and Genomic Medicine*; 2022 jun [citado 4 de julio de 2023]. Disponible en: <http://medrxiv.org/lookup/doi/10.1101/2022.06.20.22275744>
36. Varela AR, Florez LJH, Tamayo-Cabeza G, Contreras-Arrieta S, Restrepo SR, Laajaj R, et al. Factors Associated With SARS-CoV-2 Infection in Bogotá, Colombia: Results From a Large Epidemiological Surveillance Study. *Lancet Reg Health - Am*. octubre de 2021;2:100048.
37. CDC. Centers for Disease Control and Prevention. 2020 [citado 20 de diciembre de 2022]. COVID-19 and Your Health. Disponible en: <https://www.cdc.gov/coronavirus/2019-ncov/your-health/understanding-risk.html>
38. Angulo-Aguado M, Corredor-Orlandelli D, Carrillo-Martínez JC, Gonzalez-Cornejo M, Pineda-Mateus E, Rojas C, et al. Association Between the LZTFL1 rs11385942 Polymorphism and COVID-19 Severity in Colombian Population. *Front Med*. 20 de junio de 2022;9:910098.
39. Li MY, Li L, Zhang Y, Wang XS. Expression of the SARS-CoV-2 cell receptor gene ACE2 in a wide variety of human tissues. *Infect Dis Poverty*. diciembre de 2020;9(1):45.
40. Singh SP, Pritam M, Pandey B, Yadav TP. Microstructure, pathophysiology, and potential therapeutics of COVID-19: A comprehensive review. *J Med Virol*. enero de 2021;93(1):275-99.
41. Marik PE, Iglesias J, Varon J, Kory P. A scoping review of the pathophysiology of COVID-19. *Int J Immunopathol Pharmacol*. enero de 2021;35:205873842110480.
42. Blanco-Melo D, Nilsson-Payant BE, Liu WC, Uhl S, Hoagland D, Møller R, et al. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell*. mayo de 2020;181(5):1036-1045.e9.
43. Farag NS, Breitinger U, Breitinger HG, El Azizi MA. Viroporins and inflammasomes: A key to understand virus-induced inflammation. *Int J Biochem Cell Biol*. mayo de 2020;122:105738.
44. Van Damme W, Dahake R, van de Pas R, Vanham G, Assefa Y. COVID-19: Does the infectious inoculum dose-response relationship contribute to understanding heterogeneity in disease severity and transmission dynamics? *Med Hypotheses*. enero de 2021;146:110431.
45. Hou YJ, Okuda K, Edwards CE, Martinez DR, Asakura T, Dinnon KH, et al. SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. *Cell*. julio de 2020;182(2):429-446.e14.
46. Leisman DE, Deutschman CS, Legrand M. Facing COVID-19 in the ICU: vascular dysfunction, thrombosis, and dysregulated inflammation. *Intensive Care Med*. junio de 2020;46(6):1105-8.
47. Gustine JN, Jones D. Immunopathology of Hyperinflammation in COVID-19. *Am J Pathol*. enero de 2021;191(1):4-17.
48. Guan W jie, Ni Z yi, Hu Y, Liang W hua, Ou C quan, He J xing, et al. Clinical

Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med*. 30 de abril de 2020;382(18):1708-20.

49. Wu Y, Kang L, Guo Z, Liu J, Liu M, Liang W. Incubation Period of COVID-19 Caused by Unique SARS-CoV-2 Strains: A Systematic Review and Meta-analysis. *JAMA Netw Open*. 22 de agosto de 2022;5(8):e2228008.

50. CDC. Centers for Disease Control and Prevention. 2020 [citado 20 de diciembre de 2022]. Healthcare Workers. Disponible en: <https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-care/clinical-considerations-course.html>

51. Stokes EK, Zambrano LD, Anderson KN, Marder EP, Raz KM, El Burai Felix S, et al. Coronavirus Disease 2019 Case Surveillance — United States, January 22–May 30, 2020. *MMWR Morb Mortal Wkly Rep*. 19 de junio de 2020;69(24):759-65.

52. Whitaker M, Elliott J, Bodinier B, Barclay W, Ward H, Cooke G, et al. Variant-specific symptoms of COVID-19 in a study of 1,542,510 adults in England. *Nat Commun*. 11 de noviembre de 2022;13(1):6856.

53. Ma Q, Liu J, Liu Q, Kang L, Liu R, Jing W, et al. Global Percentage of Asymptomatic SARS-CoV-2 Infections Among the Tested Population and Individuals With Confirmed COVID-19 Diagnosis: A Systematic Review and Meta-analysis. *JAMA Netw Open*. 14 de diciembre de 2021;4(12):e2137257.

54. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA*. 7 de abril de 2020;323(13):1239.

55. Saavedra Trujillo CH. Consenso colombiano de atención, diagnóstico y manejo de la infección por SARS-COV-2/COVID 19 en establecimientos de atención de la salud. Recomendaciones basadas en consenso de expertos e informadas en la evidencia. *Infectio*. 26 de marzo de 2020;24(3):1.

56. Coronavirus disease (COVID-19): Post COVID-19 condition [Internet]. 2022 [citado 20 de diciembre de 2022]. Coronavirus disease (COVID-19): Post COVID-19 condition. Disponible en: [https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-\(covid-19\)-post-covid-19-condition](https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-(covid-19)-post-covid-19-condition)

57. Nasserie T, Hittle M, Goodman SN. Assessment of the Frequency and Variety of Persistent Symptoms Among Patients With COVID-19: A Systematic Review. *JAMA Netw Open*. 26 de mayo de 2021;4(5):e2111417.

58. Carvalho-Schneider C, Laurent E, Lemaigen A, Beaufils E, Bourbao-Tournois C, Laribi S, et al. Follow-up of adults with noncritical COVID-19 two months after symptom onset. *Clin Microbiol Infect*. febrero de 2021;27(2):258-63.

59. Arnold DT, Hamilton FW, Milne A, Morley AJ, Viner J, Attwood M, et al. Patient outcomes after hospitalisation with COVID-19 and implications for follow-up: results from a prospective UK cohort. *Thorax*. abril de 2021;76(4):399-401.

60. Moreno-Pérez O, Merino E, Leon-Ramirez JM, Andres M, Ramos JM, Arenas-Jiménez J, et al. Post-acute COVID-19 syndrome. Incidence and risk factors: A Mediterranean cohort study. *J Infect*. marzo de 2021;82(3):378-83.

61. Huang C, Huang L, Wang Y, Li X, Ren L, Gu X, et al. 6-month consequences

of COVID-19 in patients discharged from hospital: a cohort study. *The Lancet*. enero de 2021;397(10270):220-32.

62. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, et al. Post-acute COVID-19 syndrome. *Nat Med*. abril de 2021;27(4):601-15.

63. Brodin P, Casari G, Townsend L, O'Farrelly C, Tancevski I, Löffler-Ragg J, et al. Studying severe long COVID to understand post-infectious disorders beyond COVID-19. *Nat Med*. mayo de 2022;28(5):879-82.

64. Crook H, Raza S, Nowell J, Young M, Edison P. Long covid—mechanisms, risk factors, and management. *BMJ*. 26 de julio de 2021;n1648.

65. Sudre CH, Murray B, Varsavsky T, Graham MS, Penfold RS, Bowyer RC, et al. Attributes and predictors of long COVID. *Nat Med*. abril de 2021;27(4):626-31.

66. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, Sepulveda R, Rebolledo PA, Cuapio A, et al. More than 50 Long-term effects of COVID-19: a systematic review and meta-analysis [Internet]. *Infectious Diseases (except HIV/AIDS)*; 2021 ene [citado 20 de diciembre de 2022]. Disponible en:

<http://medrxiv.org/lookup/doi/10.1101/2021.01.27.21250617>

67. Imam Z, Odish F, Gill I, O'Connor D, Armstrong J, Vanood A, et al. Older age and comorbidity are independent mortality predictors in a large cohort of 1305 COVID-19 patients in Michigan, United States. *J Intern Med*. octubre de 2020;288(4):469-76.

68. Yanez ND, Weiss NS, Romand JA, Treggiari MM. COVID-19 mortality risk for older men and women. *BMC Public Health*. diciembre de 2020;20(1):1742.

69. Yang J, Zheng Y, Gou X, Pu K, Chen Z, Guo Q, et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *Int J Infect Dis*. mayo de 2020;94:91-5.

70. Fadini GP, Morieri ML, Longato E, Avogaro A. Prevalence and impact of diabetes among people infected with SARS-CoV-2. *J Endocrinol Invest*. junio de 2020;43(6):867-9.

71. Khan MMA, Khan MN, Mustagir MG, Rana J, Islam MS, Kabir MI. Effects of underlying morbidities on the occurrence of deaths in COVID-19 patients: A systematic review and meta-analysis. *J Glob Health*. diciembre de 2020;10(2):020503.

72. Zhou Y, Yang Q, Chi J, Dong B, Lv W, Shen L, et al. Comorbidities and the risk of severe or fatal outcomes associated with coronavirus disease 2019: A systematic review and meta-analysis. *Int J Infect Dis*. octubre de 2020;99:47-56.

73. Fond G, Nemani K, Etchecopar-Etchart D, Loundou A, Goff DC, Lee SW, et al. Association Between Mental Health Disorders and Mortality Among Patients With COVID-19 in 7 Countries: A Systematic Review and Meta-analysis. *JAMA Psychiatry*. 1 de noviembre de 2021;78(11):1208.

74. Liu N, Sun J, Wang X, Zhao M, Huang Q, Li H. The Impact of Dementia on the Clinical Outcome of COVID-19: A Systematic Review and Meta-Analysis. *J Alzheimers Dis*. 8 de diciembre de 2020;78(4):1775-82.

75. Yang J, Hu J, Zhu C. Obesity aggravates COVID-19: A systematic review and meta-analysis. *J Med Virol*. enero de 2021;93(1):257-61.

76. Patanavanich R, Glantz SA. Smoking Is Associated With COVID-19 Progression: A Meta-analysis. *Nicotine Tob Res.* 24 de agosto de 2020;22(9):1653-6.
77. Western Cape Department of Health in collaboration with the National Institute for Communicable Diseases, South Africa. Erratum to: Risk Factors for Coronavirus Disease 2019 (COVID-19) Death in a Population Cohort Study from the Western Cape Province, South Africa. *Clin Infect Dis.* 9 de abril de 2022;74(7):1321-1321.
78. Bhaskaran K, Rentsch CT, MacKenna B, Schultze A, Mehrkar A, Bates CJ, et al. HIV infection and COVID-19 death: a population-based cohort analysis of UK primary care data and linked national death registrations within the OpenSAFELY platform. *Lancet HIV.* enero de 2021;8(1):e24-32.
79. Ssentongo P, Ssentongo AE, Heilbrunn ES, Ba DM, Chinchilli VM. Association of cardiovascular disease and 10 other pre-existing comorbidities with COVID-19 mortality: A systematic review and meta-analysis. Hirst JA, editor. *PLOS ONE.* 26 de agosto de 2020;15(8):e0238215.
80. Brenner EJ, Ungaro RC, Geary RB, Kaplan GG, Kissous-Hunt M, Lewis JD, et al. Corticosteroids, But Not TNF Antagonists, Are Associated With Adverse COVID-19 Outcomes in Patients With Inflammatory Bowel Diseases: Results From an International Registry. *Gastroenterology.* agosto de 2020;159(2):481-491.e3.
81. Michelena X, Borrell H, López-Corbeto M, López-Lasanta M, Moreno E, Pascual-Pastor M, et al. Incidence of COVID-19 in a cohort of adult and paediatric patients with rheumatic diseases treated with targeted biologic and synthetic disease-modifying anti-rheumatic drugs. *Semin Arthritis Rheum.* agosto de 2020;50(4):564-70.
82. CMMID COVID-19 Working Group, Davies NG, Jarvis CI, Edmunds WJ, Jewell NP, Diaz-Ordaz K, et al. Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature.* 13 de mayo de 2021;593(7858):270-4.
83. The COVID-19 Genomics UK (COG-UK) consortium, Volz E, Mishra S, Chand M, Barrett JC, Johnson R, et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature.* 13 de mayo de 2021;593(7858):266-9.
84. Uriu K, Kimura I, Shirakawa K, Takaori-Kondo A, Nakada T aki, Kaneda A, et al. Neutralization of the SARS-CoV-2 Mu Variant by Convalescent and Vaccine Serum. *N Engl J Med.* 16 de diciembre de 2021;385(25):2397-9.
85. Alvarez-Moreno C, Valderrama-Beltran SL, Silva R, De La Hoz Siegler IH, Varón Vega FA, Martínez-Vernaza S, et al. Associated factors for mortality in a COVID-19 colombian cohort: is the third wave relevant when Mu variant was predominant epidemiologically? *Travel Med Infect Dis.* mayo de 2023;53:102579.
86. Karlsen TH. Understanding COVID-19 through genome-wide association studies. *Nat Genet.* abril de 2022;54(4):368-9.
87. de Araújo JLF, Menezes D, Aguiar RS de, Souza RP de. IFITM3, FURIN, ACE1, and TNF- α Genetic Association With COVID-19 Outcomes: Systematic Review and Meta-Analysis. *Front Genet.* 1 de abril de 2022;13:775246.
88. Downes DJ, Cross AR, Hua P, Roberts N, Schwessinger R, Cutler AJ, et al.

Identification of LZTFL1 as a candidate effector gene at a COVID-19 risk locus. *Nat Genet.* noviembre de 2021;53(11):1606-15.

89. Espinas NA, Tu LN, Furci L, Shimajiri Y, Harukawa Y, Miura S, et al. Transcriptional regulation of genes bearing intronic heterochromatin in the rice genome. Mittelsten Scheid O, editor. *PLOS Genet.* 18 de marzo de 2020;16(3):e1008637.
90. Rose AB. Introns as Gene Regulators: A Brick on the Accelerator. *Front Genet.* 7 de febrero de 2019;9:672.
91. López-Rodríguez R, Del Pozo-Valero M, Corton M, Minguez P, Ruiz-Hornillos J, Pérez-Tomás ME, et al. Presence of rare potential pathogenic variants in subjects under 65 years old with very severe or fatal COVID-19. *Sci Rep.* 20 de junio de 2022;12(1):10369.
92. Kosmicki JA, Horowitz JE, Banerjee N, Lanche R, Marcketta A, Maxwell E, et al. A catalog of associations between rare coding variants and COVID-19 outcomes [Internet]. *Infectious Diseases (except HIV/AIDS)*; 2020 nov [citado 20 de diciembre de 2022]. Disponible en: <http://medrxiv.org/lookup/doi/10.1101/2020.10.28.20221804>
93. Butler-Laporte G, Povysil G, Kosmicki JA, Cirulli ET, Drivas T, Furini S, et al. Exome-wide association study to identify rare variants influencing COVID-19 outcomes: Results from the Host Genetics Initiative [Internet]. *Infectious Diseases (except HIV/AIDS)*; 2022 mar [citado 20 de diciembre de 2022]. Disponible en: <http://medrxiv.org/lookup/doi/10.1101/2022.03.28.22273040>
94. Bats ML, Rucheton B, Fleur T, Orioux A, Chemin C, Rubin S, et al. Covichem: A biochemical severity risk score of COVID-19 upon hospital admission. Lazzeri C, editor. *PLOS ONE.* 6 de mayo de 2021;16(5):e0250956.
95. Haimovich AD, Ravindra NG, Stoytchev S, Young HP, Wilson FP, van Dijk D, et al. Development and Validation of the Quick COVID-19 Severity Index: A Prognostic Tool for Early Clinical Decompensation. *Ann Emerg Med.* octubre de 2020;76(4):442-53.
96. Dite GS, Murphy NM, Spaeth E, Allman R, Lifelines Corona Research Initiative. Validation of a clinical and genetic model for predicting severe COVID-19. *Epidemiol Infect.* 2022;150:e91.
97. MGI sequencing platforms: High-throughput gene sequencers, DNBSEQ™ sequencing technology-MGI-Leading Life Science Innovation [Internet]. [citado 29 de diciembre de 2022]. Disponible en: <https://en.mgi-tech.com/products/>
98. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 15 de agosto de 2009;25(16):2078-9.
99. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet.* octubre de 2016;99(4):877-85.
100. CDC. Centers for Disease Control and Prevention. 2020 [citado 20 de diciembre de 2022]. Coronavirus Disease 2019 (COVID-19). Disponible en: <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html>

101. Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng J, et al. Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. *J Infect.* agosto de 2020;81(2):e16-25.
102. Lim S, Bae JH, Kwon HS, Nauck MA. COVID-19 and diabetes mellitus: from pathophysiology to clinical management. *Nat Rev Endocrinol.* enero de 2021;17(1):11-30.
103. Liu D, Zhang T, Wang Y, Xia L. The Centrality of Obesity in the Course of Severe COVID-19. *Front Endocrinol.* 11 de marzo de 2021;12:620566.
104. He X, Cheng X, Feng X, Wan H, Chen S, Xiong M. Clinical Symptom Differences Between Mild and Severe COVID-19 Patients in China: A Meta-Analysis. *Front Public Health.* 14 de enero de 2021;8:561264.
105. Parasher A. COVID-19: Current understanding of its Pathophysiology, Clinical presentation and Treatment. *Postgrad Med J.* mayo de 2021;97(1147):312-20.
106. Foster KJ, Jauregui E, Tajudeen B, Bishehsari F, Mahdavinia M. Smell loss is a prognostic factor for lower severity of coronavirus disease 2019. *Ann Allergy Asthma Immunol.* octubre de 2020;125(4):481-3.
107. Peghin M, Palese A, Venturini M, De Martino M, Gerussi V, Graziano E, et al. Post-COVID-19 symptoms 6 months after acute infection among hospitalized and non-hospitalized patients. *Clin Microbiol Infect.* octubre de 2021;27(10):1507-13.
108. Asadi-Pooya AA, Akbari A, Emami A, Lotfi M, Rostamihosseinkhani M, Nemati H, et al. Risk Factors Associated with Long COVID Syndrome: A Retrospective Study. *Iran J Med Sci [Internet].* noviembre de 2021 [citado 20 de diciembre de 2022];46(6). Disponible en: <https://doi.org/10.30476/ijms.2021.92080.2326>
109. de Oliveira JF, de Ávila RE, de Oliveira NR, da Cunha Severino Sampaio N, Botelho M, Gonçalves FA, et al. Persistent symptoms, quality of life, and risk factors in long COVID: a cross-sectional study of hospitalized patients in Brazil. *Int J Infect Dis.* septiembre de 2022;122:1044-51.
110. Bordon Y. A new set of genes linked to critical COVID-19. *Nat Rev Immunol.* abril de 2022;22(4):208-208.
111. Zheng J, Zhang Y, Zhao H, Liu Y, Baird D, Karim MA, et al. Multi-ancestry Mendelian randomization of omics traits revealing drug targets of COVID-19 severity. *eBioMedicine.* julio de 2022;81:104112.
112. Rao S, Baranova A, Cao H, Chen J, Zhang X, Zhang F. Genetic mechanisms of COVID-19 and its association with smoking and alcohol consumption. *Brief Bioinform.* 5 de noviembre de 2021;22(6):bbab284.
113. Zhang HL, Wu J, Zhu J. The Immune-Modulatory Role of Apolipoprotein E with Emphasis on Multiple Sclerosis and Experimental Autoimmune Encephalomyelitis. *Clin Dev Immunol.* 2010;2010:1-10.
114. Ostendorf BN, Patel MA, Bilanovic J, Hoffmann HH, Carrasco SE, Rice CM, et al. Common human genetic variants of APOE impact murine COVID-19 mortality. *Nature.* 10 de noviembre de 2022;611(7935):346-51.
115. van der Made CI, Netea MG, van der Veerdonk FL, Hoischen A. Clinical implications of host genetic variation and susceptibility to severe or critical COVID-

19. *Genome Med.* 19 de agosto de 2022;14(1):96.
116. van der Made CI, Simons A, Schuurs-Hoeijmakers J, van den Heuvel G, Mantere T, Kersten S, et al. Presence of Genetic Variants Among Young Men With Severe COVID-19. *JAMA.* 18 de agosto de 2020;324(7):663.
117. Stikker BS, Stik G, van Ouwkerk AF, Trap L, Spicuglia S, Hendriks RW, et al. Severe COVID-19-associated variants linked to chemokine receptor gene control in monocytes and macrophages. *Genome Biol.* diciembre de 2022;23(1):96.
118. Khalil BA, Noha Mousaad Elemam, Maghazachi AA. Chemokines and chemokine receptors during COVID-19 infection. *Comput Struct Biotechnol J.* 2021;19:976-88.
119. Voloudakis G, Vicari JM, Venkatesh S, Hoffman GE, Dobrindt K, Zhang W, et al. A translational genomics approach identifies IL10RB as the top candidate gene target for COVID-19 susceptibility. *Npj Genomic Med.* 5 de septiembre de 2022;7(1):52.
120. Fricke-Galindo I, Falfán-Valencia R. Genetics Insight for COVID-19 Susceptibility and Severity: A Review. *Front Immunol.* 1 de abril de 2021;12:622176.
121. Chen Y, Zhou X, Yan H, Huang H, Li S, Jiang Z, et al. CANPT Score: A Tool to Predict Severe COVID-19 on Admission. *Front Med.* 18 de febrero de 2021;8:608107.
122. Woolhouse M. How to make predictions about future infectious disease risks. *Philos Trans R Soc B Biol Sci.* 12 de julio de 2011;366(1573):2045-54.
123. Mergenthaler C, Gurp M, Rood E, Bakker M. The study of spatial autocorrelation for infectious disease epidemiology decision-making: a systematized literature review. *CABI Rev.* 15 de julio de 2022;2022:cabireviews202217018.