



**Universidad de Concepción
Facultad Farmacia**

**GENOTIPIFICACIÓN DEL POLIMORFISMO DE LA
ENZIMA CONVERTIDORA DE ANGIOTENSINA 2
rs2285666 Y DE LA PROTEASA
TRANSMEMBRANA SERINA 2 rs12329760
ASOCIADOS A ENFERMEDAD COVID 19 EN
POBLACIÓN DE LA REGIÓN DEL BÍO BÍO
POR MYCHEL ANDREA URRUTIA BURGOS**

Trabajo de Investigación de Internado en Sector Salud presentado en la
Facultad de Farmacia de la Universidad de Concepción para optar al título
profesional de Bioquímico

Profesora Guía y Patrocinante

Dra. Liliana Ivone Lamperti Fernández
Departamento de Bioquímica Clínica e Inmunología
Facultad de Farmacia
Universidad de Concepción

Mayo, 2023

Concepción, Chile

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.

AGRADECIMIENTOS

El proceso fue más complejo de lo que pensé en algún momento. Sin embargo, si hubiese sido fácil, no tendría todas las herramientas y conocimientos que tengo ahora. No hablo sólo de conocimientos académicos, cascadas de señalización, valores de referencia e interpretación de casos clínicos, también hablo de perseverancia, persistencia, capacidad de superación y confianza. Confianza que muchas veces se perdió y me costó volver a encontrar, o que probablemente fue aumentando en el camino, ya que basalmente se encontraba en concentraciones muy bajas.

Quiero agradecer en primera instancia a mi familia. A mi padre, Bernardo, por enseñarme sus valores y ser mi inspiración. A mi abuela, Mireya, por darme todo su amor y dedicación durante todos mis procesos. A mi tata, Heberto, por ser la persona más bondadosa y esforzada que he conocido. A mi madre, Claudia, por darme su apoyo cuando lo necesité. A mi hermano, Mateo, por darme su alegría. A mi pareja, Juan José, por acompañarme en este proceso y por siempre confiar en mí. A mi querida abuela paterna que me acompaña desde el cielo. Sé lo orgullosa que hubiese estado.

A mis profesores que me acompañaron todos estos años. A los profesores del Depto. de Bioquímica Clínica e Inmunología, a mi profesora Guía y patrocinante, la Dra. Liliana Lamperti, a quien agradezco sinceramente por darme la oportunidad de trabajar en este proyecto. A mis profesores de ramos clínicos, a quienes les tengo una gran admiración. Especialmente a quien recuerdo con mucho cariño, el Dr. Daniel Durán (Q.E.P.D), quien fue parte importante de mi aprendizaje en el área clínica. Al Instituto de Salud Pública, por permitirme realizar este análisis que aporta con nuevos conocimientos en población chilena. Al Laboratorio Clínico PreveGen, por su constante apoyo y respaldo.

A mis amigos y compañeros, quienes fueron pacientes y comprensivos cuando los planes no se podían llevar a cabo por los estudios. A mis amigas STEM, quienes me hubiese gustado conocer antes, pero que en este poco tiempo les he tomado mucho cariño. En especial a Javiera Appel, por ser mi compañera durante toda la carrera, por siempre confiar en mis capacidades y ser mi confidente.

Muchas veces admiramos tanto a otras personas que queremos ser como ellos. Después de altos y bajos, descubrí que quiero ser yo misma.

TABLA DE CONTENIDOS

AGRADECIMIENTOS	iii
TABLA DE CONTENIDOS	v
ÍNDICE DE TABLAS	viii
ÍNDICE DE FIGURAS.....	x
RESUMEN.....	xii
ABSTRACT	xiv
1. INTRODUCCIÓN	17
2. PROBLEMA DE INVESTIGACIÓN.....	31
3. PREGUNTA DE INVESTIGACIÓN	32
4. HIPÓTESIS	33
5. OBJETIVO GENERAL.....	34
6. OBJETIVOS ESPECÍFICOS.....	35
7. METODOLOGÍA	36
7.1. Diseño del estudio	36
7.2. Tamaño muestral	39
7.3. Selección y obtención de la muestra.....	43

7.4. Extracción de ADN genómico.....	45
7.5. Cuantificación de ADN	47
7.6. Electroforesis en gel de agarosa 1%.....	48
7.7. Estandarización y ensayo de genotipificación de SNPs	49
7.8. Interpretación de resultados	52
7.9. Análisis estadístico	54
8. RESULTADOS.....	57
8.1. Caracterización de la población	57
8.2. Análisis de los extractos de ADN.	61
8.3. Estandarización de qPCR	65
8.4. Análisis de curvas de qPCR	71
8.5. Determinación de distribución y frecuencia genotípica y alélica.....	74
8.6. Asociación entre variables clínicas, patologías y genotipos	78
9. DISCUSIÓN	85
10. CONCLUSIÓN Y PROYECCIONES.....	95
11. GLOSARIO	97
12. BIBLIOGRAFÍA	99
13. ANEXOS.....	111

13.1. Anexo 1. Resumen proyecto “Respuesta Inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile” y consentimiento informado.	111
13.2. Anexo 2. Documento de aprobación del Comité Ética Científico Concepción.	116
13.3. Anexo 3. Cálculo tamaño muestral para el proyecto “Respuesta inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile”.....	118
13.4. Anexo 4. Formulario de datos autoreportados por los voluntarios del proyecto “Respuesta inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile”.....	120
13.5. Anexo 5. Clasificación y manejo de residuos biológicos.	122

ÍNDICE DE TABLAS

Tabla 1. Criterios de inclusión y exclusión para el universo muestral del proyecto: “Respuesta inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile”.	38
Tabla 2. Indicadores demográficos, epidemiológicos y de vacunación para regiones seleccionadas.	40
Tabla 3. Tamaño muestral esperado según origen de la muestra para el proyecto del Instituto de Salud Pública.....	42
Tabla 4. Tamaño muestral total, regional esperado y regional real según origen de muestra del proyecto del Instituto de Salud.	42
Tabla 5. Secuencia contexto del gen ACE2 y TMPRSS2, en donde se encuentran los polimorfismos y fue diseñada la sonda	50
Tabla 6. Características de población de estudio según clasificación y variables.	58
Tabla 7. Concentración (ng/μL) y relación A260/A280 de las muestras extraídas	62

Tabla 8. Datos obtenidos desde curva estándar para polimorfismo rs2285666	68
Tabla 9. Datos obtenidos desde curva estándar para polimorfismo rs12329760	70
Tabla 10. Genotipos obtenidos del total de muestras (n=98).	73
Tabla 11. Cálculo de frecuencias alélicas observadas y esperadas para ambos polimorfismos.....	77
Tabla 12. Genotipo y análisis estadístico de los genotipos vs patologías	79

ÍNDICE DE FIGURAS

Figura 1. Expresión de la enzima convertidora de angiotensina 2 (ACE2) en diferentes tejidos.....	20
Figura 2. Entrada del virus SARS-CoV-2 mediante el receptor ACE2 humano	21
Figura 3. Representación esquemática del gen de ACE2 y de su estructura tridimensional.....	25
Figura 4. Esquema representativo de posibles curvas de amplificación para determinación de genotipos en las muestras.....	53
Figura 5. Caracterización de la población de estudio según clasificación y variables clínicas.....	60
Figura 6. Gráficos de concentración de muestras y relación 260/280.....	62
Figura 7. Electroforesis en gel de agarosa para extractos de ADN.....	64
Figura 8. Curvas de amplificación para curva estándar polimorfismo rs2285666	67
Figura 9. Gráfico para curva estándar rs2285666.....	68

Figura 10. Curvas de amplificación para curva estándar polimorfismo rs12329760.....	69
Figura 11. Curva estándar para rs12329760	70
Figura 12. Curvas de amplificación para muestras rs2285666	72
Figura 13. Distribución genotípica de las muestras de la población de estudio.	75
Figura 14. Distribución alélica de las muestras de la población de estudio	76
Figura 15. Análisis de componentes principales (PCA) para los polimorfismos.	82
Figura 16. Análisis de método de SHAP para los polimorfismos	83
Figura 17. Beeswarm plot para los polimorfismos.....	84

RESUMEN

El síndrome respiratorio agudo severo causado por Coronavirus (SARS-CoV-2), surgió en China en diciembre de 2019 y actualmente se ha extendido por todo el mundo causando la pandemia por COVID-19 con más de 6,8 millones de muertes a nivel mundial. La morbilidad y mortalidad causadas por COVID-19 aumentan drásticamente con la edad y condiciones de salud coexistentes, incluyendo el cáncer, diabetes y las enfermedades cardiovasculares. Además de estos predisponentes, los factores genéticos tienen un papel importante en la infección por COVID-19, ya que pueden contribuir a la transmisibilidad del SARS-CoV-2 y a la progresión de la enfermedad, pero estos factores son en gran parte desconocidos. Hoy en día se conoce que la proteína ACE2 (enzima convertidora de angiotensina 2) es el receptor de SARS-CoV-2 en células humanas y en conjunto con la proteasa serina tipo 2 (TMPRSS2), son claves en la unión e internalización del virus a través de la proteína viral en la superficie de SARS-CoV-2 (proteína S o Spike). En el gen ACE2 se han identificado polimorfismos implicados en la predisposición de hipertensión arterial y actualmente, se han identificado algunos relacionados con la gravedad de la enfermedad COVID-19, como lo es el polimorfismo rs2285666, que provoca un cambio de base nucleotídica (C<T). La proteína TMPRSS2 (proteasa transmembrana serina 2) es esencial en el clivaje de la glicoproteína S del SARS-CoV-2 para su entrada al huésped y del mismo modo presenta polimorfismos asociados a la enfermedad,

como lo es el polimorfismo rs12329760 que genera un cambio, de igual manera que el anterior, de una citosina por una timina.

Es por lo anterior que el objetivo propuesto para este trabajo fue determinar la presencia del polimorfismo rs2285666 y rs12329760 en una población de la región del Biobío y su asociación con obesidad, hipertensión o diabetes. La metodología para responder a este objetivo fue en primera instancia, reclutar participantes que respondieron una encuesta de salud de enfermedades crónicas no transmisibles a los cuales se les tomó una muestra de sangre para su posterior análisis genético. Para la identificación de los polimorfismos, se aisló el ADN genómico de estas muestras y con la utilización de sondas de hidrólisis fluorescentes para PCR en tiempo real, se realizó la detección de los polimorfismos rs2285666 y rs12329760 y su posible asociación con diabetes, hipertensión u obesidad.

Los resultados mostraron la presencia del alelo de riesgo para ambos polimorfismos, con una distribución genotípica para rs2285666 de 63,27% (homocigoto *wild type*, CC), 18,37% (heterocigoto, CT) y 18,37% (homocigoto mutado, TT) y para rs12329760 de 88,78% (homocigoto *wild type*) y 11,22% (heterocigoto mutado, CT). En relación a las frecuencias alélicas de la población, se encontró un 27,55% del alelo de riesgo (T) para rs2285666 y un 5,61% del

alelo de riesgo (T) para rs12329760. En cuanto a la asociación de estos polimorfismos a patologías reportadas, no se encontraron diferencias estadísticamente significativas.

ABSTRACT

The severe acute respiratory syndrome caused by Coronavirus (SARS-CoV-2), emerged in China in December 2019 and has currently spread throughout the world causing the COVID-19 pandemic with more than 6.8 million deaths worldwide. Morbidity and mortality caused by COVID-19 increased with age and coexisting health conditions, including cancer, diabetes, and cardiovascular disease. In addition to these predisposing factors, genetic factors play an important role in COVID-19 infection, as they may contribute to SARS-CoV-2 transmissibility and disease progression, but these factors are largely unknown. Today it is known that the ACE2 protein (angiotensin-converting enzyme 2) is the receptor for SARS-CoV-2 in human cells and, together with the type 2 serine protease (TMPRSS2), are key in the binding and internalization of the virus, through the viral protein on the surface of SARS-CoV-2 (S or Spike protein). In the ACE2 gene, polymorphisms involved in the predisposition to arterial hypertension have been identified and, currently, some have been identified related to the severity of the COVID-19 disease, such as the rs2285666 polymorphism, which causes a nucleotide base change (C< T) and is located in

intron 3 of this gene. The TMPRSS2 protein (transmembrane protease serine 2) is essential in the cleavage of the S glycoprotein S of SARS-CoV-2 for its entry into the host and likewise presents polymorphisms associated with the disease, such as the rs12329760 polymorphism that generates a change , in the same way as the previous one, from a cytosine to a thymine, a change that is generated in exon 6 of this gene.

For this reason, the objective proposed for this work was to determine the presence of the rs2285666 and rs12329760 polymorphism in a population from the Biobío region and to determine the prevalence of risk alleles for COVID-19 infection in this population. Also, it was determined if these polymorphisms had any association with different pathologies reported by the volunteers (obesity, hypertension and/or diabetes). The methodology to respond to this objective was, in the first instance, to recruit participants who answered a health survey on chronic non-communicable diseases, from whom a blood sample was taken for subsequent genetic analysis. To identify the polymorphisms, the genomic DNA of these samples was isolated and with the use of fluorescent hydrolysis probes for real-time PCR, the detection of the rs2285666 and rs12329760 polymorphisms and their possible association with diabetes, hypertension or obesity was carried out.

The results showed the presence of the risk allele for both polymorphisms, with a genotypic distribution for rs2285666 of 63.27% (homozygous wild type, CC), 18.37% (heterozygous, CT) and 18.37% (homozygous mutated, TT) and for rs12329760 of 88.78% (wild type homozygous) and 11.22% (mutated heterozygous, CT). In relation to the allele frequencies of the population, 27.55% of the risk allele (T) was found for rs2285666 and 5.61% of the risk allele (T) for rs12329760. Regarding the association of these polymorphisms with reported pathologies, no statistically significant differences were found.

1. INTRODUCCIÓN

En diciembre de 2019, un grupo de pacientes con neumonía de causa desconocida se vinculó a un mercado mayorista de productos del mar en Wuhan, China. Se descubrió un betacoronavirus previamente desconocido mediante el uso de secuenciación imparcial en muestras de estos pacientes [1]. Se confirmaba posteriormente que este nuevo betacoronavirus era parte de la gran familia de los coronavirus (CoVs). Los CoVs están ampliamente distribuidos en la naturaleza y representan una seria amenaza para los humanos y una variedad de huéspedes mamíferos, causando enfermedades respiratorias, gastrointestinales y del sistema nervioso central. Los CoVs son virus de ARN monocatenario de sentido positivo no segmentados con envoltura y se clasifican en α -, β -, γ - y δ [2]. Más recientemente, tres β -coronavirus causantes del Síndrome Respiratorio Agudo Severo (SARS-CoV) en 2003 [1], coronavirus causante del Síndrome Respiratorio del Medio Oriente en 2012 (MERS-CoV) [1-3] y SARS-CoV-2 en 2019, han cruzado la barrera de las especies para infectar a los humanos resultando en enfermedades respiratorias, incluida la neumonía, que pueden ser fatales. Sin duda, este último es el que más daño ha causado a nivel mundial, provocando la actual pandemia denominada COVID-19 (Corona Virus Disease 2019, por sus siglas en inglés) y que ha provocado más de 6,9 millones de muertes y 776 millones de casos a nivel mundial a la fecha [4].

Las manifestaciones clínicas de la enfermedad causada por SARS-CoV-2 varían de compromiso leve a graves. Los síntomas informados más comunes son fiebre, tos, mialgia o fatiga, seguidos de neumonía y disnea, mientras que los síntomas notificados menos comunes incluyen dolor de cabeza, diarrea y hemoptisis. Se informó que los pacientes con síntomas leves se recuperaron después de una semana, mientras que los casos graves experimentaron insuficiencia respiratoria progresiva debido al daño alveolar, que probablemente condujo a la muerte [5]. Además de las manifestaciones clínicas, se ha identificado que el modo principal de transmisión de persona a persona es a través de gotitas de secreción de mucus respiratorio y contacto directo. La transmisión por gotitas ocurre cuando una persona habla, estornuda o tose, y el virus se libera de las secreciones respiratorias. Al entrar en contacto directo con la membrana mucosa de un paciente infectado, las gotitas tienden a transmitir el virus. Las gotas no viajan más de un metro y medio y no permanecen en el aire. Esto genera incertidumbre con respecto al mecanismo de transmisión. Quizás existen otras formas posibles por las cuales una persona puede infectarse, por ejemplo, tocando superficies u objetos que tienen el virus y luego tocándose la boca, la nariz o los ojos [6]. Sin embargo, esta es sólo una de las intrigas que nos deja la pandemia del COVID-19, junto con el origen, los factores predisponentes, factores de gravedad y mutaciones del virus.

Lo que sí tiene claro la comunidad científica el día de hoy, es como se replica el virus SARS-CoV-2. Esta comienza cuando la glicoproteína S viral en la superficie del virus se une a su receptor que es la enzima convertidora de angiotensina 2 (ACE2, angiotensin converting enzyme 2, por sus siglas en inglés) en la célula huésped [7]. El receptor ACE2 se expresa en células epiteliales dentro de una variedad de órganos que incluyen pulmones, riñones y vasos sanguíneos [8]. Además, se sabe que existe una alta expresión de ACE2 en células alveolares tipo II (ACT2) de pulmón [9], células epiteliales superiores y estratificadas del esófago, enterocitos del íleon y colon [10], colangiocitos [11], células miocárdicas, células del túbulo proximal del riñón y células uroteliales de la vejiga [9] como se muestra en la figura 1, además de otros órganos. Estos hallazgos indicaron que aquellos órganos con un alto nivel de células que expresan ACE2 podrían considerarse como de alto riesgo potencial de infección por SARS-CoV-2 [9].

Al igual que SARS-CoV, SARS-CoV-2 utiliza el mismo receptor ACE2 para ingresar e infectar una célula huésped. Esta característica se debe a la gran similitud en las relaciones filogenéticas y genómicas entre ambos coronavirus [12]. La secuencia genética que codifica para la proteína S de SARS-CoV-2 y de SARS-CoV, tiene una similitud de alrededor de un 78%, y para el dominio RBD, una similitud de un 76% [13].

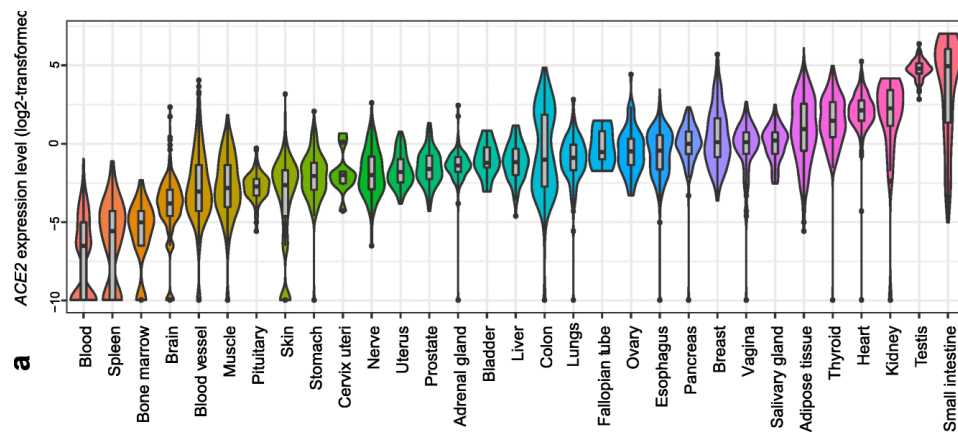


Figura 1. Expresión de la enzima convertidora de angiotensina 2 (ACE2) en diferentes tejidos. [14]

La proteína S de SARS-CoV-2 posee dos subunidades, S1 y S2. En la subunidad S1 se localiza el dominio RBD, el cual, permite el reconocimiento y la unión al receptor específico de la célula huésped, mientras que, la subunidad S2 permite la fusión de la membrana viral con la membrana celular [13]. Después de la exitosa unión de la proteína S al receptor ACE2, esta es procesada proteolíticamente por una serina proteasa de transmembrana de tipo II (TMPRSS2) [15]. TMPRSS2 produce una escisión en dos diferentes posiciones del dominio S2, lo que permite, la separación del dominio RBD, la activación de la proteína S y la posterior fusión de ambas membranas. De esta manera, la partícula vírica ingresa en la célula huésped y se produce la liberación de su material genético en el citoplasma de la célula infectada [16]. como se muestra en la figura 2.

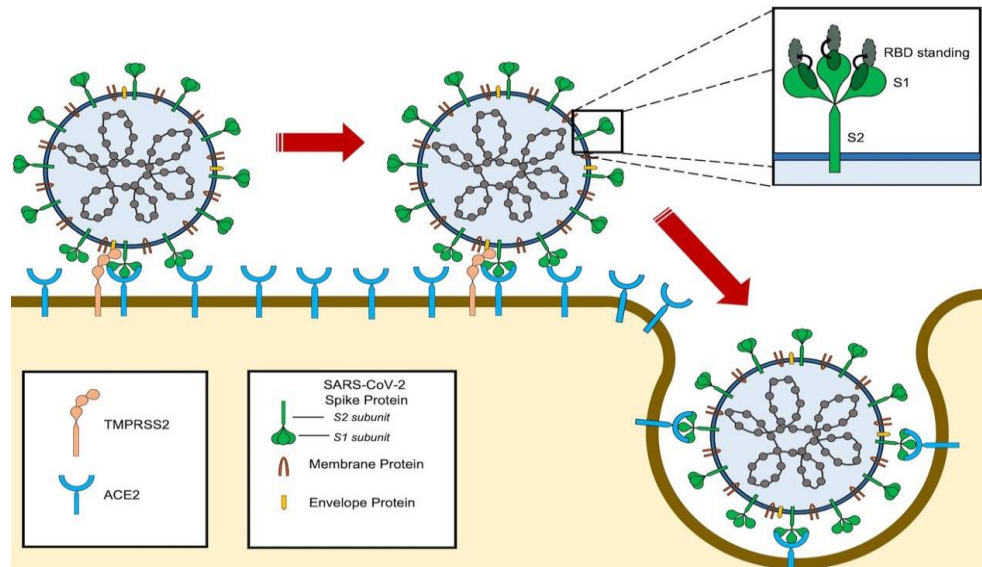


Figura 2. Entrada del virus SARS-CoV-2 mediante el receptor ACE2 humano [17].

Describiendo al receptor del virus SARS-CoV-2, la proteína ACE2 es una aminopeptidasa asociada a la membrana plasmática que fue descubierta en el año 2000 como un homólogo de la enzima convertidora de la angiotensina (ACE) [18]. Se ha visto que esta enzima posee una potente regulación negativa del sistema renina-angiotensina-aldosterona (RAAS), en donde el RAAS juega un papel importante en la fisiología cardiovascular y en la homeostasis a través de la regulación del equilibrio electrolítico, la presión arterial y el tono vascular [19]. La ACE cataliza la formación de angiotensina II (Ang II, un potente

vasoconstrictor) a partir de angiotensina I, desempeñando un papel clave en el control de la función cardiorenal y en el control de la presión arterial [16]. Una disfunción del RAAS resulta en concentraciones elevadas de Ang II, lo que contribuye al aumento del estrés oxidativo, la inflamación y, además, al desarrollo de síndrome metabólico, diabetes y sus complicaciones asociadas, como retinopatía diabética [20]. Por otro lado, ACE2 cataliza la formación de angiotensina (1-7) (Ang 1-7) a partir de Ang II, lo que permite la reducción de la inflamación y del estrés oxidativo estimulado por Ang II, ya que, Ang 1-7 posee efectos opuestos a Ang II, tales como, vasodilatación, efecto antiinflamatorio y efecto antiproliferativo. También, cataliza la formación de angiotensina (1-9) a partir de angiotensina 1 (Ang I), que posee características antihipertensivas y cardioprotectoras. [19, 21]. Es muy importante destacar que la homeostasis del RAAS es crítica para la regulación fisiológica y patológica de varios órganos, incluidos el corazón, los riñones y los pulmones [19].

El gen de ACE2 se encuentra localizado en la sub-banda 2, banda 2, región 2 del brazo p en el cromosoma X (Xp22.2) [22]. La secuencia de la proteína ACE2 comprenden 805 residuos de aminoácidos y contiene una secuencia consenso HEXXH de unión a Zn^{2+} (donde X es cualquier aminoácido con carga) [23]. Esta enzima consta de una región extracelular constituida por dos dominios, el primero localizado en la región N-terminal que corresponde al dominio carboxipeptidasa

(PD: peptidase domain; residuos del 19 al 611) y el segundo dominio denominado de colectrina (CLD: collectrin domain; del residuo 612 al 740) (Fig. 3A). La región transmembrana (TM) abarca los residuos 741-761; y por último la parte intracelular de ACE2 que corresponde a los residuos 762-805 dentro de la región C-terminal. Es importante mencionar que el sitio de unión del virus SARS-CoV-2 se encuentra en el dominio carboxipeptidasa (PD) (Fig. 3B). La ACE2 también existe en forma soluble, ya que puede ser escindida en la región PC (sitio de escisión de peptidasas). Aunque no se conoce bien su función, una característica que mantiene a pesar de que no se encuentra anclada a la membrana es que no pierde su dominio PD activo, esto es de gran relevancia porque es el sitio de interacción con el virus SARS-CoV-2, y esto podría ser de ayuda para disminuir la interacción del virus con ACE2 de membrana y reducir la infección en las células [24].

Existe una serie de estudios que vinculan la variación de alelos en el gen de ACE2 con múltiples enfermedades, en donde la mayoría son polimorfismos. Un polimorfismo implica una de dos o más variantes de una secuencia particular de ADN. El tipo más común de polimorfismo implica la variación en un solo par de bases, los llamados polimorfismos de nucleótido sencillo, o SNP que están siendo estudiados por los científicos para ver su correlación en el genoma humano con enfermedades, respuesta a los fármacos, y otros fenotipos [26]. Un

polimorfismo del gen ACE2 se documentó por primera vez en la población china con tres variantes de ACE2 (rs4240157, rs4646155 y rs4830542) asociadas con HTA (hipertensión arterial) [27]. En una cohorte canadiense de adolescentes con dependencia de nicotina; rs2074192, rs233575 y rs2158083 se asociaron significativamente con variaciones patológicas de la presión arterial [28]. En la India, el estudio de 246 pacientes con HTA y 274 personas normotensas indicó una asociación de HTA con la mutación ACE2 rs21068809 [29]. La vía RAAS también puede estar regulada por un polimorfismo en la ACE2. En afroamericanos con hipertensión se informó un polimorfismo de ACE2 [30].

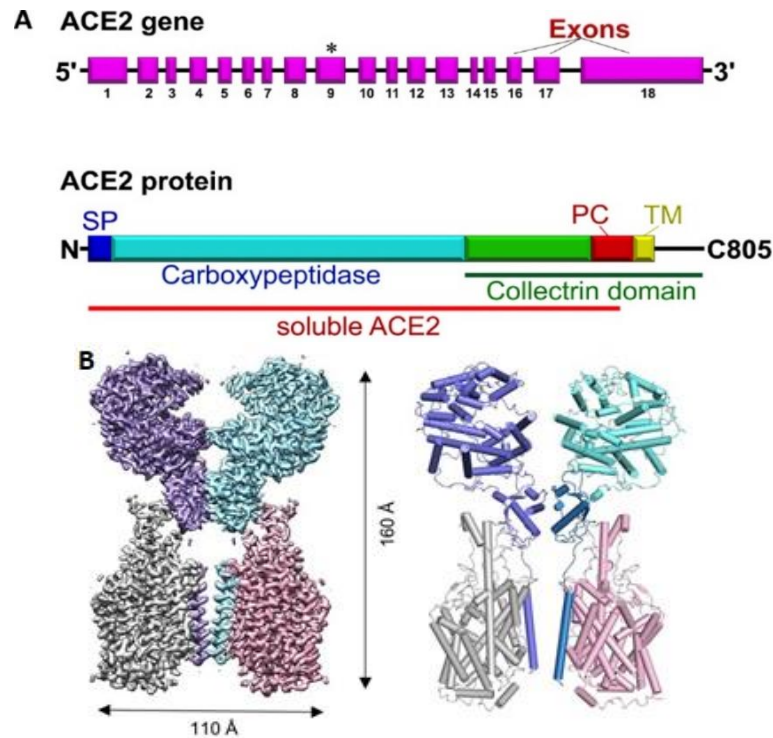


Figura 3. Representación esquemática del gen de ACE2 y de su estructura tridimensional. A) Representación esquemática del gen ACE2. El asterisco indica la codificación del sitio de unión al zinc (HEXXH). SP: péptido señal. PC: sitio de escisión de peptidasa. TM: dominio transmembrana [25]. B) Estructura de ACE2 en su forma dimérica (las estructuras en morado y celeste corresponden cada una a un monómero) obtenida por criomicroscopía electrónica con una alta resolución (2.9 Å). ACE2 está formando un complejo con el transportador de aminoácidos neutros dependiente de sodio (B°AT1; color gris y rosa) [24].

Hoy en día existen estudios que postulan polimorfismos del gen ACE2 que podrían aumentar o bien disminuir la afinidad con la región RBD de la proteína S del SARS-CoV-2. Es posible que polimorfismos del gen ACE2, influyan en la susceptibilidad al SARS-CoV-2 y en el resultado de la enfermedad COVID-19 [31]. Se han evaluado ciertas variantes (SNPs) de ACE2 que puedan alterar la susceptibilidad del huésped al SARS-CoV-2 al afectar esta interacción. Suryamohan y colaboradores analizaron más de 290.000 muestras que representan a más de 400 grupos de población de conjuntos de datos genómicos públicos e identificaron múltiples variantes de proteínas ACE2 determinantes en la unión con la proteína S de SARS-CoV-2. Utilizando los datos estructurales informados, se identificaron variantes de ACE2 naturales que podrían afectar potencialmente la interacción virus-huésped y, por lo tanto, alterar la susceptibilidad del huésped. Estas incluyen las variantes S19P, I21V, E23K, K26R, T27A, N64K, T92I, Q102P y H378R que se predijo que aumentarían la susceptibilidad, mientras que las variantes K31R, N33I, H34R, E35K, E37K, D38V, Y50F, N51S, M62V, K68E, F72V Se predijo que Y83H, G326E, G352V, D355N, Q388L y D509Y eran variantes protectoras que muestran una unión disminuida a la proteína S. Usando ensayos bioquímicos, se confirmó que K31R y E37K tenían afinidad disminuida, y las variantes K26R y T92I mostraron una mayor afinidad por la proteína S en comparación con la ACE2 WT (wild type). De acuerdo con esto, ACE2 K26R y T92I solubles resultaron más efectivos para

bloquear la entrada de un pseudovirus, lo que sugiere que las variantes de ACE2 pueden modular la susceptibilidad al SARS-CoV-2 [32].

La serina proteasa 2 transmembrana (TMPRSS2) es una proteína de la superficie celular que se expresa principalmente en las células endoteliales de los tractos respiratorio y digestivo. Como serina proteasa, participa en la escisión de los enlaces peptídicos de proteínas que tienen serina como aminoácido nucleofílico dentro del sitio activo. La función biológica exacta de TMPRSS2 se desconoce en gran medida, aunque la investigación ha demostrado que está involucrada en ciertas patologías [33]. En cáncer de próstata, por ejemplo, se han visto translocaciones de protooncogenes como ETS, ERG o ETV1 junto con TMPRSS2, siendo la translocación observada en el 72 % de todos los cánceres de próstata según la investigación actual [34]. El gen que codifica a esta proteína se ubica en el cromosoma 21 (21q22.3) cuya actividad transcripcional está controlada por receptores de andrógenos [33].

Sin duda, los polimorfismos asociados a enfermedades ayudan a entender de otra manera la fisiopatología de ellas. Estudios vinculan aquellos que puedan tener alguna relación con la sintomatología grave de la enfermedad COVID-19. Por ejemplo, un estudio realizado en población egipcia de 147 controles sanos y 299 pacientes COVID-19 genotipificados con los alelos de riesgo para los

polimorfismos rs2285666 y rs12329760, fueron significativamente vinculados con la gravedad de la enfermedad de COVID-19, al igual que ciertas comorbilidades (hipertensión, diabetes) y parámetros de laboratorio. Ambos SNP se encontraban entre los predictores más altos de gravedad de la enfermedad: TMPRSS2 rs12329760 CT + TT [odds ratio (IC 95 %) 17,6 (5,1–61,10), ACE2 rs2285666 CT + TT 9,9 (3,2-30,9), ambos $p < 0,001$]. Hubo un aumento en la expresión de frecuencias genotípicas de ACE2 rs2285666 y TMPRSS2 rs1232976 (TT), (CT + TT) y (T) alelo en el grupo de COVID-19 grave en comparación con los grupos de control y leve. La gravedad de la enfermedad también se relacionó con niveles elevados de proteína C reactiva (PCR), ferritina y dímero D, y niveles más bajos de linfocitos y recuento de plaquetas (todos $p < 0,001$) [35].

El SNP rs2285666 es una variante ubicada en el cromosoma Xp22 en el intrón 3. Se sugiere que esta variante puede alterar el empalme del ARNm y, por lo tanto, afectar la expresión del gen ACE2 [36]. Varios estudios han demostrado asociaciones con hipertensión, enfermedad coronaria y diabetes combinada con accidente cerebrovascular [36, 37]. Curiosamente, el alelo C se asocia con un aumento de los niveles séricos de ACE2 en controles sanos, sujetos con diabetes y pacientes con diabetes junto con accidente cerebrovascular [37]. Aunque se desconoce si el aumento de los niveles séricos de ACE2 asociados con el alelo C en realidad también refleja un aumento de ACE2 unido a la membrana en, por

ejemplo, el pulmón u otros tejidos, este polimorfismo parece ser un candidato atractivo asociado con una susceptibilidad alterada a la infección. o curso de COVID-19 dado el hecho de que ACE2 funciona como un 'receptor' celular para la entrada del virus.

Para el SNP rs12329760, que sustituye una citosina por una timina, se ha identificado como variante con sentido erróneo, que provoca la sustitución de valina por metionina en la proteína TMPRSS2 en la posición 160. El análisis computacional ha demostrado que la variante rs12329760 puede afectar la estructura de esta enzima al formar un bolsillo en la estructura de la proteína. [38] Aunque el análisis computacional sobre la influencia de la variante rs12329760 en la función de la enzima TMPRSS2 [38] predijo que teóricamente podría aumentar la carga viral al reducir la entrada viral a las células huésped, el vínculo real entre la carga viral y la gravedad de la enfermedad aún es pobremente entendido [39]. Por ejemplo, una investigación mostró una asociación entre esta variante y la carga viral a pesar de la falta de asociación entre esta variante y la gravedad de la enfermedad [40]. Por lo tanto, puede que no haya una relación simple entre la carga viral y la gravedad de la enfermedad, y el mecanismo molecular detrás del papel de rs12329760 en la gravedad de la enfermedad, debe aclararse en futuras investigaciones.

Finalmente, dado que el virus se ha transmitido en humanos a partir de murciélagos y/u otros huéspedes intermediarios, se ha realizado un trabajo considerable para explorar la secuencia y las variaciones estructurales en las proteínas S de SARS-CoV-2 y los receptores ACE2 y TMPRSS2 del huésped [8]. Se ha demostrado que las variaciones en ambos (La proteína S viral y las secuencias del huésped) pueden actuar como una barrera para la infección viral en todas las especies [41, 42]. Esto, a su vez, plantea dos posibilidades interesantes: primero, ¿el polimorfismo genético natural en los genes del huésped y/o la proteína ACE2 humana influye en la afinidad con la proteína S del SARS-CoV-2? En segundo lugar, ¿son estas variantes genéticas benignas o podrían considerarse como una compensación evolutiva en la que una variación en un rasgo puede traer ventajas y desventajas a la aptitud de un organismo? Este trabajo de internado tiene como objetivo en primer lugar explorar si existen las variantes rs2285666 y rs12329760 en población chilena y si existen diferencias en su prevalencia de acuerdo a variables metabólicas. Los hallazgos brindarían información interesante sobre la posible relación entre las variantes genéticas en la ACE2 y TMPRSS2 humana y la susceptibilidad y/o resistencia contra la infección por COVID-19.

2. PROBLEMA DE INVESTIGACIÓN

La pandemia por COVID-19 causado por el virus SARS-CoV-2, se ha convertido en el problema más amenazante para todas las poblaciones del mundo de las últimas décadas. Es importante recalcar que los factores genéticos poseen un papel importante en diferentes enfermedades, siendo estos de gran ayuda para entender la fisiopatología y tratamientos, pero aún son en gran parte desconocidos para esta nueva enfermedad. Se conoce al receptor ACE2 y la proteína TMPRSS2 como los determinantes en la entrada del virus y para ambos se han reportado polimorfismos que se asocian a severidad del cuadro clínico de COVID-19. Sin embargo, no se sabe de esta asociación ni su prevalencia en población chilena.

3. PREGUNTA DE INVESTIGACIÓN

En base a lo anterior, se formula la siguiente pregunta de investigación:

¿Cuál es la prevalencia de los alelos de riesgo del polimorfismo rs2285666 (ACE2) y rs12329760 (TMPRSS2) en población chilena de la región del Biobío y su asociación con datos demográficos y de enfermedades crónicas?

4. HIPÓTESIS

El alelo de riesgo para la infección por el virus SARS-CoV-2 de los polimorfismos rs2285666 (ACE2) y rs12329760 (TMPRSS2) está presente en población de la región del Biobío y se puede relacionar con obesidad, hipertensión o diabetes.

5. OBJETIVO GENERAL

Determinar la presencia del polimorfismo rs2285666 y rs12329760 en una población de la región del Biobío y su asociación con obesidad, hipertensión o diabetes.

.

6. OBJETIVOS ESPECÍFICOS

1. Generar los criterios de reclutamiento y selección de una muestra en población chilena.
2. Extraer el ADN genómico según criterios de amplificación para los ensayos de genotipificación de SNPs.
3. Realizar la estandarización del ensayo de genotipificación de los polimorfismos rs2285666 y rs12329760.
4. Analizar estadísticamente la prevalencia de los alelos de los polimorfismos rs2285666 y rs12329760 y su asociación con obesidad, hipertensión o diabetes.

7. METODOLOGÍA

7.1. Diseño del estudio

Este trabajo de investigación en Internado en Sector Salud, está enmarcado dentro de un proyecto del Instituto de Salud Pública denominado: “Respuesta inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile” (Anexo 1). Este último tuvo como objetivo general caracterizar y analizar la respuesta inmune estimulada por la infección de SARS-CoV-2 y por la vacunación contra SARS-CoV-2. Se ejecutó como un estudio observacional transversal, realizado en dos regiones de Chile: Metropolitana y del Biobío. La población a estudiar estuvo compuesta por hombres y mujeres, entre 18 y 80 años de edad con antecedentes de infección por SARS CoV-2; y por hombres y mujeres en el mismo grupo etario, vacunados contra SARS CoV-2. El estudio se desarrolló en 24 semanas, periodo durante el que se obtuvieron las muestras necesarias para la evaluación inmunológica humoral y celular. Particularmente, en la Región del Biobío participó CESFAM Víctor Manuel Fernández (CVMF) de la comuna de Concepción como unidad de toma de muestras y el laboratorio del Departamento de Bioquímica Clínica e Inmunología de la Facultad de Farmacia de la Universidad de Concepción, funcionó como centro de operaciones para la separación y envío de muestras al Instituto de Salud Pública, en donde se

realizaron los análisis específicos para cumplir con los objetivos propios de ese proyecto.

Paralelamente a este estudio, se incorporó la evaluación genética de los polimorfismos rs2285666 y rs12329760, lo que fue sometido a revisión y aprobación por el Comité de Ética Científico del Servicio de Salud Concepción (Anexo 2) y que fue el objetivo principal de este trabajo de investigación de Internado en Sector Salud.

Los criterios de inclusión y exclusión para el proyecto del Instituto de Salud Pública, se resumen en la tabla 1, dando a conocer las características del universo muestral que se obtuvo. Los criterios fueron divididos en dos grupos:

- (i) Post-infectados: personas recuperadas de COVID 19
- (ii) Vacunados: personas con esquema de vacunación completa contra SARS-CoV-2 para vacuna Pfizer o Sinovac.

Tabla 1. Criterios de inclusión y exclusión para el universo muestral del proyecto: “Respuesta inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile”.

(i) Post-infectados		(ii) Vacunados	
Inclusión	Exclusión	Inclusión	Exclusión
- Hombres y mujeres mayores o iguales a 18 años y hasta 80 años.	- Hombres y mujeres menores de 18 años y mayores de 80 años.	- Hombres y mujeres mayores o iguales a 18 años y hasta 80 años.	- Hombres y mujeres menores de 18 años y mayores de 80 años.
- Con RT-PCR negativo a SARS-CoV-2 al momento toma muestra	- Contraindicación médica para la donación de sangre.	- Con RT-PCR negativo a SARS-CoV-2 al momento toma muestra	- Haber completado el esquema de inmunización para vacuna Pfizer o Sinovac en menos de 21 días de la toma de esta muestra.
- Personas recuperadas por COVID-19 con un tiempo máximo de 8 meses post-infección, y que hayan requerido una hospitalización.	- Individuos recuperados de COVID-19, pero que hayan recibido alguna de las vacunas contra SARS-CoV-2.	- Personas con un esquema de vacunación completo para vacuna Pfizer o Sinovac con un mínimo de 21 días y no superior a 6 meses. Se considera esquema de vacunación completo al recibir las dos dosis de la misma vacuna.	- Contraindicación médica para la donación de sangre.
- Firma de consentimiento informado. Anexo 1.	- Personas con antecedentes de reinfección por SARS-CoV-2 (dos periodos diferentes de infección por SARS-CoV-2).	- Personas sin antecedentes de infección previa por SARS-CoV-2	- Individuos que estén vacunados contra SARS-CoV-2 pero que hayan sido previamente infectados por SARS-CoV-2
	- Sintomatología asociada a COVID-19, como: ▪ Fiebre. ▪ Tos seca ▪ Cansancio ▪ Dolor de garganta ▪ Diarrea ▪ Conjuntivitis ▪ Dolor de cabeza ▪ Pérdida del sentido del olfato o del gusto ▪ Erupciones cutáneas o pérdida del color en los dedos de las manos o de los pies ▪ Dificultad para respirar o sensación de falta de aire ▪ Dolor o presión en el pecho ▪ Incapacidad para hablar o moverse	- Firma de consentimiento informado. Anexo 1	- Sintomatología asociada a COVID-19, como: ▪ Fiebre. ▪ Tos seca ▪ Cansancio ▪ Dolor de garganta ▪ Diarrea ▪ Conjuntivitis ▪ Dolor de cabeza ▪ Pérdida del sentido del olfato o del gusto ▪ Erupciones cutáneas o pérdida del color en los dedos de las manos o de los pies ▪ Dificultad para respirar o sensación de falta de aire ▪ Dolor o presión en el pecho ▪ Incapacidad para hablar o moverse

7.2. Tamaño muestral

El Instituto de Salud Pública realizó una caracterización poblacional y epidemiológica del desarrollo de la pandemia en las regiones a estudiar (Metropolitana y Biobío) con el fin de utilizar estos datos para poder estimar la proporción de individuos que alcanzaron inmunogenicidad esperada en los dos grupos: los post-infectados (no vacunados) y vacunados, considerando las dos vacunas que se estaban implementando en población chilena hasta esa fecha. Los datos que se obtuvieron con fecha 7 de abril de 2021 se describen en la tabla 2.

Tabla 2. Indicadores demográficos, epidemiológicos y de vacunación para regiones seleccionadas.

<i>Indicador</i>	<i>Metropolitana</i>	<i>Biobío</i>
<i>Información Demográfica¹</i>		
<i>Hombre</i>	3.462.267	750.730
<i>Mujer</i>	3.650.541	806.075
<i>Total Población</i>	7.112.808	1.556.805
<i>0 a 19 años</i>	1.870.650	429.746
<i>20 a 59 años</i>	4.146.257	865.386
<i>60 o más años</i>	1.095.901	261.673
<i>Información vigilancia COVID-19²</i>		
<i>Total Casos acumulados</i>	506.210	119.273
<i>Fallecidos totales</i>	16.875	2.175
<i>Tasa Incidencia acumulada</i>	6.230,2	7.169,2
<i>Información vacunación³</i>		
<i>Población objetivo</i>	6.430.040	1.290.996
<i>Total Pfizer</i>	360.823	123.919
<i>Total Sinovac</i>	4.041.444	938.087
<i>Total Vacunas</i>	4.402.267	1.062.006
<i>% personas con al menos 1 dosis</i>	43,9%	51,3%
<i>% personas con 1 y 2 dosis</i>	25,4%	31,1%

1: Censo de Población y Vivienda, INE 2017; 2: Informe Epidemiológico N°109. Departamento de Epidemiología, 22-03-2021; 4: Plan Nacional de Vacunación COVID-19, 7-04-2021.

Con estos datos, se calculó el tamaño mínimo muestral para que el estimador, y sus respectivos intervalos de confianza, representen la inmunogenicidad en: post-infectados, vacunados con CoronaVac y vacunados con Pfizer-BioNTech, utilizando la ecuación de cálculo de tamaño muestral desarrollada en el Anexo 3.

El tamaño de muestra total y esperado corresponde a 749 pacientes desglosados según indica la tabla 3. El tamaño muestral esperado para cada región fue de un 50% del total y ajustable según origen de la muestra, en especial para el caso de NO vacunado (post -infección). Sin embargo, producto de la diferencia de vacunación en la población (mayoritariamente vacunados con Coronavac) y la alta tasa de vacunación que tenía la población del Bíobio en esa fecha, se obtuvo el tamaño muestral real para la Región del Bíobio indicado en la tabla 4:

Tabla 3. Tamaño muestral esperado según origen de la muestra para el proyecto del Instituto de Salud Pública.

Origen de la Muestra	Tamaño muestral total	Región del BíoBío
Vacunados con Pfizer-BioNTech	223	111
Vacunados con Coronavac	254	127
No vacunados (post-infectados)	272	136
Total	749	374

Tabla 4. Tamaño muestral total, regional esperado y regional real según origen de muestra del proyecto del Instituto de Salud.

Origen de la Muestra	Tamaño muestral total	Región del BíoBío (esperado)	Región del BíoBío (real)
Vacunados con Pfizer-BioNTech	223	111	56
Vacunados con Coronavac	254	127	135
No vacunados (post-infectados)	272	136	5
Total	749	374	196

7.3. Selección y obtención de la muestra

Los participantes fueron llamados y reclutados según las fuentes de datos con los registros de cada participante en el establecimiento (CVMF). Los datos de aquellas personas que cumplieron con los criterios, así como las respuestas del formulario presentada en el Anexo 4, fueron registrados por EU. Pablo López y EU. Blanca Quijada del CESFAM Víctor Manuel Fernández y subidos en la Plataforma de Estudios del Instituto de Salud Pública de Chile. Para mantener la confidencialidad de los datos de los participantes, las muestras fueron anonimizadas y se les asignó un número serial interno en la unidad de toma de muestras. Las personas seleccionadas fueron convocadas un día específico (definido por cada coordinador de establecimiento en conjunto con coordinador del estudio) y solo participaron en el estudio aquellas que accedan a firmar un consentimiento informado (Anexo 1). Posterior a la firma de este consentimiento y llenado del formulario, se procedió a la toma de una muestra de hisopado nasofaríngeo (HNF) y cuatro muestras de sangre de 5 mL cada una. Cabe destacar que, para efectos de este trabajo de investigación, sólo se utilizó un tubo de sangre al vacío con anticoagulante etilendiaminotetracético EDTA (5 mL) con el fin de obtener la fracción de leucocitos para su posterior análisis genético.

Luego de obtener los documentos y las muestras de los participantes, se procedió a establecer criterios de selección para las muestras del trabajo de investigación en Internado en Sector Salud. Se realizó un muestreo aleatorio simple, seleccionando un total de 98 muestras del universo muestral, para luego clasificarlos según si eran sanos o presentaban patología (obesidad, diabetes, hipertensión, 2 o más patologías), grupo etario y género. Lo anterior con el fin de comparar estas variables con los diferentes genotipos encontrados. Estos datos fueron obtenidos desde la encuesta que respondieron los voluntarios (Anexo 4).

Para efectos de eliminación de residuos biológicos de forma segura, cada enfermero encargado de la toma de muestras realizó la eliminación según protocolo del CESFAM. Por otro lado, el procesamiento de las muestras en la Universidad de Concepción y sus desechos biológicos, fueron clasificados según el instructivo elaborado en colaboración con la Facultad de Ciencias Biológicas (Anexo 5) y eliminados por MATPEL.

7.4. Extracción de ADN genómico

La obtención de ADN genómico se realizó desde la fracción de leucocitos, la cual se rescató desde la centrifugación de muestras de sangre periférica obtenidas en tubos al vacío con anticoagulante EDTA a 1400 x g durante 10 min. Se removió cuidadosamente el plasma y la capa de leucocitos se almacenó en tubos de 1,5 mL a -80°C hasta su uso, rotulados con su respectivo código de anonimización.

La extracción de ADN genómico se realizó desde la fracción de leucocitos, siguiendo el protocolo de micrométodo de extracción de ADN de *Salazar et al., 2001* [43]. Para comenzar con la lisis celular, se transfirieron 300 µL de fracción de leucocitos a un tubo estéril de 1,5 mL. Se adicionó igual volumen de yoduro de sodio (NaI) 6 M, luego de mezclar en vortex por 20 segundos y adicionar 600 µL cloroformo/alcohol isoamílico (v/v 24:1) para la precipitación de proteínas. Se volvió a mezclar en vortex por 20 segundos y se centrifugó a 16260 x g por 8 minutos. Posteriormente, se realizó la separación de fases y precipitación de ADN genómico, removiendo cuidadosamente la fase superior sin arrastrar la fase inferior compuesta por glóbulos rojos lisados, la cual fue transferida a otro tubo estéril de 1,5 mL. Se adicionaron 400 µL de isopropanol al 70% y se incubó a

temperatura ambiente por 3 minutos. Durante esta incubación, se agitó por inversión varias veces. Luego, se centrifugó a 16260 x g por 5 minutos y se eliminó cuidadosamente el sobrenadante mediante aspiración sin arrastrar el pellet. Luego, se lavó nuevamente el pellet con 500 µL de isopropanol al 37% y se centrifugó a 16260 x g por 5 minutos. Se removió cuidadosamente el isopropanol que estaba en la fase superior mediante aspiración y el pellet se dejó secar al aire de 5 a 8 minutos. Después, se procedió a hidratar y resuspender el ADN en 50 µL de buffer TRIS-EDTA (pH 8,0) y se agitó en vortex para homogenizar. Finalmente, se incubaron los extractos por 10 minutos a 56°C y posteriormente almacenados a -20°C.

7.5. Cuantificación de ADN

Se procedió a determinar la concentración y pureza de los extractos obtenidos por medio de cuantificación de ADN a través de lecturas de absorbancia. Para la realización de esta cuantificación, fueron tomados 2 μL de los extractos de ADN y puestos en la placa de microvolumen Take3™ dentro del lector de microplacas Epoch™ (BioTek). Como blanco se utilizaron 2 μL de buffer TE (pH 8,0). Las muestras fueron leídas a 260 nm en el equipo lector de microplacas donde se midió la concentración obtenida en cada muestra en unidades de concentración ($\text{ng}/\mu\text{L}$). y también, se realizó la lectura a 280 nm para conocer la pureza de las muestras, utilizando la relación A_{260}/A_{280} .

7.6. Electroforesis en gel de agarosa 1%

Para verificar integridad del ADN extraído y cuantificado, se preparó un gel de agarosa al 1% según el siguiente protocolo: se mezcló 1 g de agarosa en 100 mL de tampón Tris-acetato-EDTA (TAE) 1X, homogenizado en microondas por intervalos de 30 segundos hasta obtener una solución transparente sin dejar que hierva. Se dejó enfriar hasta que el matraz pueda ser palpable y se adicionaron 5 μ L de SafeView™ 1X. Luego, esta mezcla fue vertida y solidificada en un cassette horizontal con peineta formadora de carriles para el depósito y corrida de muestras durante la electroforesis. Se ubicó el gel de agarosa en cámara electroforética con 1 litro de TAE 1X. Posteriormente, se cargaron las muestras, primero el Ladder (5 μ L) y luego las muestras, previamente homogenizadas con buffer de carga 6X (muestra: se calculó el volumen necesario para cargar 1 μ g; buffer de carga: se calculó volumen para quedar en 1X; agua libre de nucleasas: se agregó hasta volumen final). Finalmente, se conectaron los electrodos a la fuente de poder con orientación de corrida hacia el polo positivo a 90 Volts por 50 minutos. Finalmente, se visualizaron las bandas electroforéticas en transiluminador de luz UV.

7.7. Estandarización y ensayo de genotipificación de SNPs

La genotipificación se realizó mediante la utilización de sondas de hidrólisis TaqMan™ de tipo MGB (*minor groove binding* o ligandos de unión al surco menor). Para la reacción en cadena de la polimerasa en tiempo real (qPCR), se utilizó el protocolo y el kit TaqMan™ SNP genotyping assay kit (Thermo Fisher Scientific, Waltham, MA, EE.UU.), número de catálogo C25516261 y C2562235320, previamente descrito para estos 2 polimorfismos (rs2285666 y rs12329760, respectivamente) [35]. La reacción de qPCR se realizó mediante el siguiente protocolo: se añadieron 7,5 µL de Master Mix TaqMan Genotyping (Applied Biosystems™), 0,75 µL de SNPs TaqMan Probes (sondas), 1 µL de ADN genómico (10 - 50 ng) y 5,75 µL de agua libre de nucleasas. El programa de qPCR estuvo compuesto por un ciclo de denaturación inicial a 95°C por 10 minutos, 40 ciclos de denaturación (95° por 15 segundos), alineamiento y extensión (60°C por 1 minuto). Este proceso fue realizado en el termociclador micPCR (BioMolecular System™), utilizando el software de la misma marca versión 2.8. La secuencia contexto en donde se diseñaron las sondas, la secuencia de los partidores y de la sonda, son presentados en la tabla X.

Tabla 5. Secuencia contexto del gen ACE2 y TMPRSS2, en donde se encuentran los polimorfismos y fue diseñada la sonda. [35]

Polimorfismo	Secuencia	Dye
rs2285666	ATAATCACTACTAAAAATTAGTAGC(C/T)TACCTGG TTCAAGTAATAAGCATTCT	VIC/FAM
rs12329760	CAGGACTTCCTCTGAGATGAGTACA(C/T)CTGAAG GATGAAGTTTGGTCCGTAG	VIC/FAM

Las sondas fueron diseñadas uniendo un fluoróforo a cada una (VIC: max A 538 nm, max E 554 nm; FAM: max A 494 nm, max E 522 nm) según el alelo con el cual va a hibridizar. Los espectros de cada fluoróforo pueden ser detectados por el termociclador en un canal distinto. Para el caso de ambos polimorfismos, el alelo *Wild Type* es el alelo C, por lo tanto, la sonda que realice el alineamiento poseerá una base complementaria T y estará ligada al fluoróforo VIC. Como se muestra en la secuencia contexto del gen de ACE2 y de TMPRSS2 (Tabla X) la detección del alelo *Wild Type* se realizó visualizando una curva de amplificación en el canal que posee el filtro de detección VIC. De la misma manera sucede con el alelo T mutado, pero en este caso la curva de amplificación se visualizó en el canal que posee el filtro de detección FAM.

La estandarización de este proceso se realizó preparando diluciones seriadas por triplicado, desde una concentración inicial de ADN (30 ng/ μ L) hasta la dilución 1/10000, para generar una curva estándar, visualizar el comportamiento de esta curva y conocer los límites de detección de la técnica. Las curvas estándar fueron realizadas para ambos polimorfismos. Luego, se comenzó a realizar las corridas de qPCR para las muestras seleccionadas en el inciso 7.3. y así, obtener su genotipo (C/C, C/T, T/T). Con estos datos se procedió a calcular la distribución alélica y genotípica, como también la frecuencia alélica y genotípica estimada basándose en el Equilibrio de Hardy Weinberg.

7.8. Interpretación de resultados

Como se explicó anteriormente, la determinación de los genotipos de las muestras se vio reflejado en las curvas de amplificación específicas para cada canal de detección. La identificación de una curva con amplificación positiva, se realizó mediante la verificación de una curva típica de amplificación de un producto de PCR con sus distintas fases: una fase inicial, una fase exponencial y una fase meseta. Los resultados fueron clasificados según los siguientes criterios y como se explica en la figura X:

- Amplificación en canal VIC, no amplificación canal FAM: genotipo CC.
- Amplificación en canal VIC, amplificación canal FAM: genotipo CT.
- No amplificación en canal VIC, amplificación en canal FAM: genotipo TT.

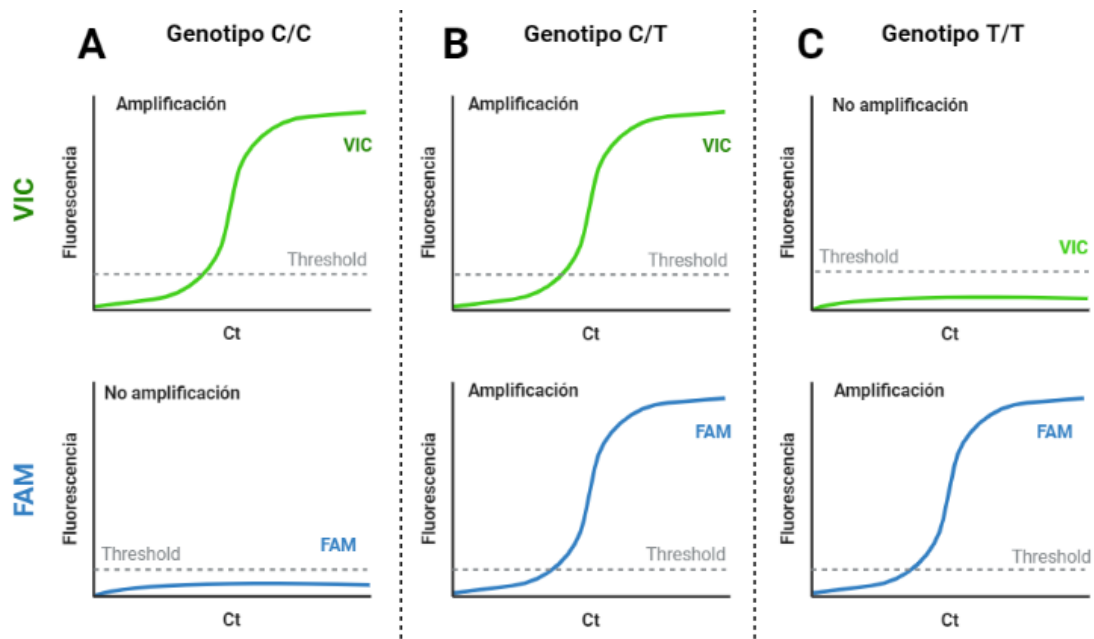


Figura 4. Esquema representativo de posibles curvas de amplificación para determinación de genotipos en las muestras. A) CC. B) CT. C) TT.

7.9. Análisis estadístico

El análisis estadístico y los gráficos fueron realizados utilizando el software GraphPad Prism 8 ®. Las variables continuas de los diferentes grupos fueron sometidas a test de normalidad mediante la prueba de D'Agostino-Pearson para $n > 5$. Para $2 < n < 5$, se utilizó la prueba de normalidad Shapiro-Wilk. Los datos que presentaron distribución normal, se expresaron como media $\pm$ desviación estándar (D.E), y se compararon con la prueba paramétrica t de Student. Por otro lado, los datos que no presentaron distribución normal se expresaron con su respectiva mediana y Rango Intercuartílico (RIC). Estos se compararon con la prueba no paramétrica U de Mann-Whitney. Para el análisis de variables de dos o más grupos, se utilizó la prueba de ANOVA para datos que presentaban distribución normal. Para los datos no paramétricos, se utilizó la prueba de Kruskal-Wallis. Las variables categóricas se expresaron como porcentaje (%) del total de datos. Se realizaron tablas de contingencia y se aplicó la prueba de Chi Cuadrado (χ^2) donde se evaluó la existencia de alguna relación entre estas variables categóricas, en cuanto a la diferencia de la proporción de individuos que presentaron alguno de los tres genotipos posibles (CC, CT, TT). Además, se evaluó la proporción de estos individuos con las siguientes patologías: obesidad;

hipertensión y diabetes, las cuales se compararon con los tres genotipos posibles.

Para el análisis multivariado de los genotipos obtenidos con las variables clínicas y las patologías, primero se realizó una reducción de dimensionalidad (reducción de las variables sin perder información) utilizando un análisis de componentes principales (PCA). Esto permitió el clustering o asociación de las variables de acuerdo a las características comunes entre ellos, de manera de descubrir las combinaciones lineales y poder explicar los patrones desconocidos entre los datos. La visualización se realizó con el método UMAP (Uniform Manifold Approximation and Project), para encontrar una representación (no lineal) de pocas dimensiones de los datos, preservando las distancias entre cada punto y sus vecinos en el espacio multidimensional. La reducción de dimensionalidad se realizó para los dos casos: ACE2 y TMPRSS2. Posteriormente, se realizó un análisis de importancia de características utilizando un modelo de clasificación utilizando herramientas de machine learning, aplicando la técnica SMOTE de sobremuestreo de minorías sintéticas para aumentar el número de casos de un conjunto de datos de forma equilibrada, de manera que el modelo de clasificación no se sobreajuste. El modelo de clasificación se implementó utilizando el clasificador Random Forest, una técnica de aprendizaje automático basado en un conjunto de árboles de decisión. Luego de generar el modelo de clasificación,

se calculó la importancia de las características en estudio con el método SHAP (Shapley Additive exPlanations), un enfoque basado en teoría de juegos, que puede explicar los resultados del modelo de clasificación, comprendiendo como impacta cada característica.

8. RESULTADOS

8.1. Caracterización de la población

La población a estudiar (n=98), fue clasificada en pacientes sanos (no reporta ninguna patología al momento de la encuesta), pacientes que reportaron obesidad, diabetes, hipertensión y si presenta 2 o más de estas patologías. Los datos se obtuvieron desde el formulario completado por cada participante (Anexo 4) Se presentan las siguientes distribuciones por grupo, según la tabla 6

Tabla 6. Características de población de estudio según clasificación y variables.

VARIABLES	Sanos	Obesidad	Hipertensión	Diabetes	2 o más
Variables clínicas	n = 48	n = 21	n = 7	n = 5	n = 17
Sexo					
Femenino	28	13	3	3	10
Masculino	20	5	5	2	7
Edad (años)	41 ± 12	36 ± 11	45 ± 13	63 ± 10	53 ± 12
Estatura (m)	1,64 ± 0,08	1,68 ± 0,11	1,65 ± 0,11	1,67 ± 0,09	1,64 ± 0,07
Peso (kg)	71 ± 10	94 ± 13	70 ± 15	70 ± 13	86 ± 15
IMC (kg/m2)	26,3 (23,9 - 28,7)	32,0 (30,5 - 33,7)	25,5 ± 3,0	25,0 ± 3,6	31,8 ± 5,2

La mayoría de las variables presentaron distribución normal, a excepción del IMC de los individuos sanos y el IMC de los individuos con obesidad. Al analizar las variables clínicas categóricas (género) y continuas (edad, estatura, peso e IMC) mediante tests estadísticos para encontrar diferencias significativas entre ellos, se encontraron los siguientes resultados mostrados en la figura 5

En cuanto al género de los voluntarios (figura 5.A), no hubo diferencias significativas en la distribución de los subgrupos (femenino y masculino) en cada grupo de clasificación. En relación a la edad (figura 5.B), se encontraron diferencias significativas entre los grupos de estudio sanos vs diabetes ($p < 0,01$), sanos vs 2 o más patologías ($p < 0,01$), obesidad vs diabetes ($p < 0,001$) y obesidad

vs 2 o más patologías ($p < 0,001$), lo que evidencia una relación entre el aumento de la edad con la aparición de patologías, hasta 2 o más. En lo que refiere al grupo de obesos y su diferencia con el grupo de diabetes y 2 o más patologías, se evidencia que, en la población estudiada la edad fue la menor de todos los grupos (promedio $\pm$ sd, 36 ± 11) y podría explicar la diferencia entre los grupos con mayor edad. Para la variable estatura (figura 5.C), no se encontraron diferencias significativas entre los grupos. Para el peso (figura 5.D), se encontraron diferencias significativas entre los grupos sanos vs obesidad ($p < 0,0001$), sanos vs 2 o más patologías ($p < 0,0001$) y obesidad con los demás grupos: hipertensión ($p < 0,001$) y diabetes ($p < 0,01$), evidenciando una correlación entre la patología obesidad con el aumento del peso. Finalmente, para el IMC (figura 5.E), se encontraron diferencias significativamente mayores ($p < 0,0001$) para los grupos sanos vs obesidad y sanos vs 2 o más patologías, apoyando lo anteriormente descrito para la variable clínica peso, ya que están estrechamente relacionadas. De igual forma, se encontraron diferencias significativamente mayores entre el grupo de los obesos vs hipertensos ($p < 0,001$), obesidad vs diabetes ($p < 0,01$) y levemente mayores para el grupo 2 o más vs hipertensos ($p < 0,05$) y 2 o más patologías vs diabetes ($p < 0,05$). Lo anterior coincide que la obesidad es una enfermedad que se define como un incremento en el IMC por sobre 29,9 y que se evidencia en los aumentos significativos con los demás grupos, que en promedio, todos poseen un IMC menor a 30, a excepción del grupo 2 o más patologías.

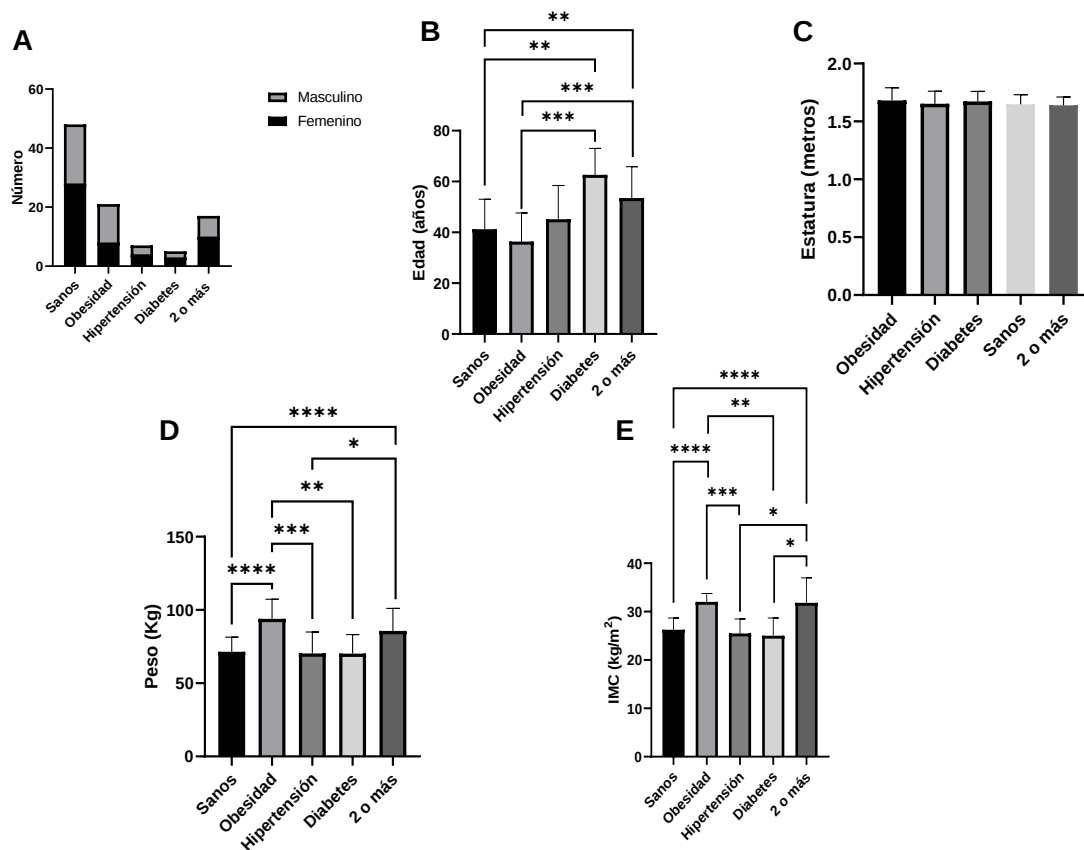


Figura 5. Caracterización de la población de estudio según clasificación y variables clínicas. A) Género B) Edad C) Estatura D) Peso E) IMC. Valores estadísticamente significativos: ns $p > 0,05$; * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$; **** $p < 0,0001$.

8.2. Análisis de los extractos de ADN.

Las muestras de sangre obtenidas y extraídas desde la fracción de leucocitos aislada de los 98 voluntarios por medio del micrométodo de extracción de ADN *Salazar et al., 2001*, fueron cuantificadas obteniéndose concentraciones de ADN entre los 16,78 y 752,9 ng/μL y cuya razón de pureza (A_{260}/A_{280}) varió entre 1,75 y 2,05, como se muestra en la tabla 7 y en la figura 6.A y 6.B.

Tabla 7. Concentración (ng/μL) y relación A₂₆₀/A₂₈₀ de las muestras extraídas.

Variables	Concentración (ng/μL)	Relación (A ₂₆₀ /A ₂₈₀)
Media	164,6	1,9
Desviación estándar	± 148,9	± 0,068
Mínimo	16,78	1,75
Máximo	752,9	2,05
Distribución normal (<i>p value</i>)	No (<0,0001)	Si (0,244)

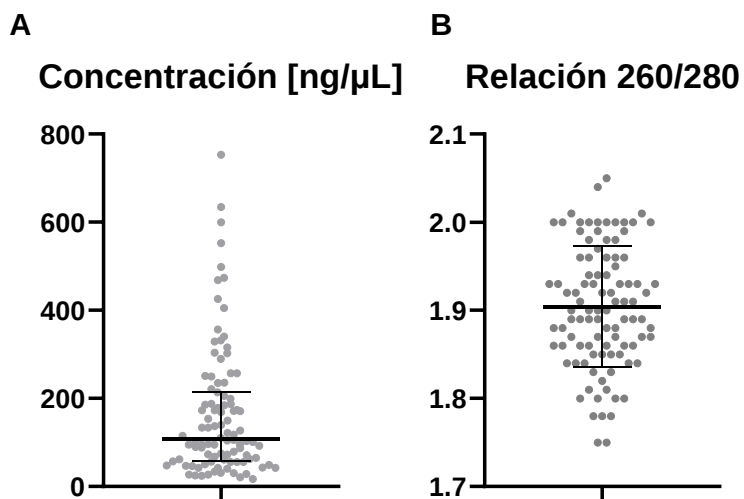


Figura 6. Gráficos de concentración de muestras y relación 260/280. A) Concentración (ng/μL). B) Relación 260/280

Se establecieron dos criterios para la aceptación o repetición de la extracción de una muestra de ADN, basándose en la cantidad mínima requerida para la realización del ensayo qPCR (1 ng/μL) y que posea un coeficiente de pureza óptimo, es decir: a. Concentración medida >15 ng/μL y b. Relación de pureza A_{260}/A_{280} entre 1,7 y 2,0. Cabe destacar, que se seleccionó una concentración >15 ng/μL por posibles repeticiones de la muestra. Como se observa en la figura 6.A, la concentración de las muestras no presentó una distribución normal según test estadísticos aplicados ($p < 0,0001$), además de presentar una gran dispersión de datos que van desde bajas hasta altas concentraciones, todas sobre 15 ng/μL. Por el contrario, la relación de pureza (A_{260}/A_{280}) (gráfico 6.B) si presentó una distribución normal ($p < 0,244$) y una baja dispersión de datos, los cuales se encontraron entre un coeficiente de 1,75 y 2,05, considerado como un rango de pureza óptima para los extractos de ADN.

Luego de la cuantificación, se procedió a analizar la integridad de los extractos con la realización de una electroforesis en gel de agarosa al 1%, de 5 muestras. Cada muestra fue elegida desde los 5 turnos de extracción que se realizaron, lo que se observa en la figura 7. Con la observación de una banda definida por encima de los 3000 pb, se evidencia y se confirma la presencia de un extracto de ADN íntegro y puro, sin degradaciones.

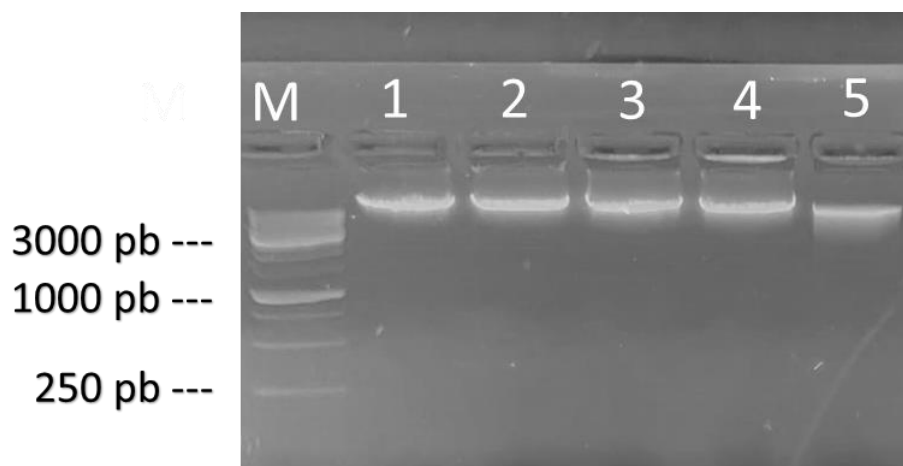


Figura 7. Electroforesis en gel de agarosa para extractos de ADN. Los extractos fueron numerados aleatoriamente desde los 5 turnos de extracción. M: marcador de pares de bases.

8.3. Estandarización de qPCR

Posterior a la verificación de integridad de los extractos de ADN, se procedió a realizar un ensayo de estandarización de la técnica de qPCR para genotipificación con TaqMan™ SNP genotyping assay kit, evaluando los siguientes criterios: 1. Visualización de una curva con forma sigmoidea, identificando sus 3 fases (inicial, exponencial y meseta). 2. Efectuar diluciones seriadas (Sin diluir 1/1 – 1/10000) de una muestra de ADN de concentración conocida, para conocer el límite de detección y los rangos de concentración de ADN óptimos de la técnica.

Se observó para el ensayo del polimorfismo rs2285666 (ACE2), una curva de amplificación característica de qPCR desde la muestra sin diluir (30 ng/μL) hasta la dilución 1/100 (Figura 8.A y 8-B). A partir de la dilución 1/1000 en adelante (Figura 8.A y 8.B, punto 4), se observa la amplificación de dos curvas. Sin embargo, por la forma de estas curvas, su baja emisión de fluorescencia y la no amplificación del triplicado completo, seguido de la no detección de una curva de amplificación para la siguiente dilución, indicó como límite de detección inferior la concentración de ADN de 0,03 ng/μL. Este mismo análisis se realizó para el polimorfismo rs12329760 (Figura 10.A y 10.B), con la diferencia de que en la

última dilución detectada (Figura 10.A y 10.B, punto 4), amplificaron los tres puntos del triplicado, sin embargo, cada curva tiene formas diferentes (inicial, exponencial y meseta) y Ct umbrales tardíos ligeramente diferentes, con baja emisión de fluorescencia.

La realización de una curva estándar, se condice con las recomendaciones del fabricante para una óptima realización del ensayo. La recomendación es utilizar rangos de concentración entre 1 – 10 ng de ADN genómico, rango en donde la curva estándar observada en las figuras 9 y 11, indican una eficiencia de reacción cercana al 80% y 90% (slope: -3,90 y -3,51; R^2 : 0,99; para rs2285666 y rs12329760, respectivamente). Adicionalmente, se verificó que se puede trabajar en un rango de concentración de ADN más amplio para la realización del ensayo, que va desde 0,03 - 30 ng/ μ L (tabla 8 y 9) y que son asociados a ciclos umbrales (Ct) con valores alrededor de los 27 y 24 (rs2285666 y rs12329760, respectivamente) para las diluciones más concentradas y para las muestras más diluidas, valores más tardíos de Ct que aparecieron en torno a los 33 y 31 (rs2285666 y rs12329760, respectivamente), los cuales se visualizaron de manera óptima antes de terminar los 40 ciclos del programa.

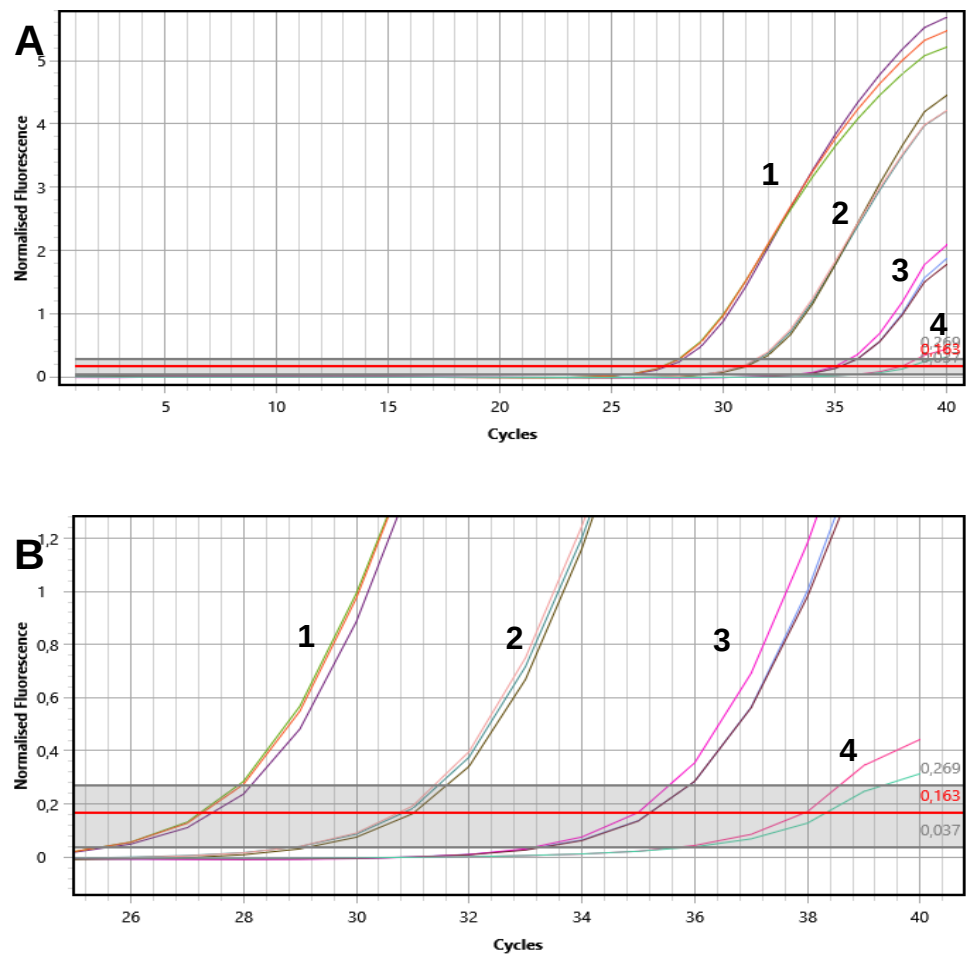


Figura 8. Curvas de amplificación para curva estándar polimorfismo rs2285666. Para A) y B) 1. Sin diluir 1/1 (30 ng/μL). 2. Dilución 1/10. 3. Dilución 1/100. 4. Dilución 1/1000.

Tabla 8. Datos obtenidos desde curva estándar para polimorfismo rs2285666.

[Dil DNA]	Log[Dil DNA]	Ct	Ct mean
1	0	27,42	27,2933333
		27,21	
		27,25	
0,1	-1	30,8	30,8366667
		31	
		30,71	
0,01	-2	35,18	35,1066667
		34,95	
		35,19	
0,001	-3	37,9	-
		38,3	

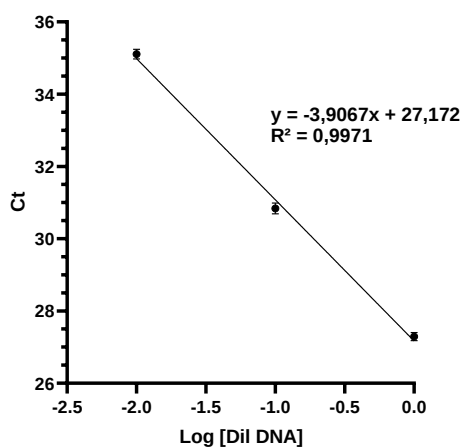


Figura 9. Gráfico para curva estándar rs2285666. Los valores fueron presentados en logaritmo de la dilución (Log [Dil DNA]) y Ct obtenido para cada dilución.

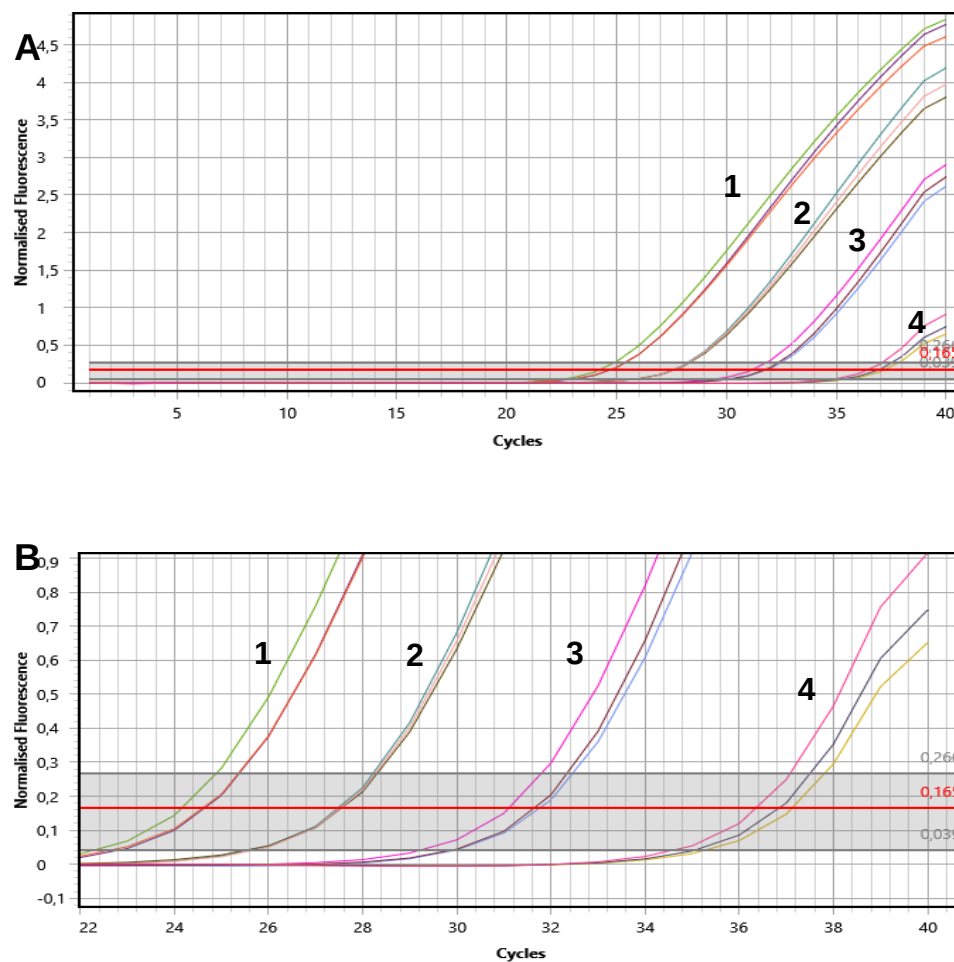


Figura 10. Curvas de amplificación para curva estándar polimorfismo rs12329760. Para A) y B) 1. Sin diluir 30 ng/uL (1/1). 2. Dilución 1/10. 3. Dilución 1/100. 4. Dilución 1/1000.

Tabla 9. Datos obtenidos desde curva estándar para polimorfismo rs12329760.

[Dil DNA]	Log[Dil DNA]	Ct	Ct mean
1	0	24,63	24,4566667
		24,15	
		24,59	
0,1	-1	27,46	27,5066667
		27,54	
		27,52	
0,01	-2	31,75	31,4933333
		31,1	
		31,63	
0,001	-3	37,12	-
		36,83	
		36,35	

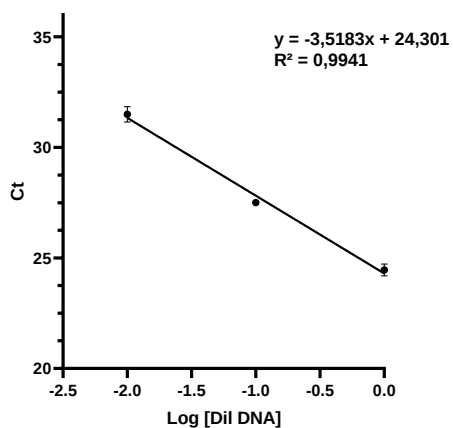


Figura 11. Curva estándar para rs12329760. Los valores fueron presentados en logaritmo de la dilución (Log [Dil DNA]) y Ct obtenido para cada dilución.

8.4. Análisis de curvas de qPCR

Luego de realizar la estandarización del ensayo de qPCR para genotipificación de muestras y conocer los rangos de concentración óptimos para su ejecución, se procedió a determinar la presencia del alelo *Wild Type* o alelo mutado para cada polimorfismo para el total de muestras obtenidas. Como se menciona en metodología, para ambos polimorfismos se detectó el alelo *Wild Type* (C) en el filtro que detecta el fluoróforo VIC (Figura 12.A) y para el alelo mutado (T), se utilizó el filtro para FAM (Figura 12.B). Se observaron curvas de amplificación para cada muestra según el alineamiento que realizó la sonda de hidrólisis TaqMan con el alelo encontrado en el gen que se presenta el polimorfismo. En la figura 12 se presenta la genotipificación de sólo 48 muestras para el polimorfismo rs2285666, ejemplificando y simplificando la modalidad de análisis y curvas de amplificación para ambos polimorfismos, debido a que esta era la capacidad máxima del equipo micPCR.

Posteriormente con estos datos, se procedió a identificar el genotipo de cada muestra (tabla 10), según lo que se indicó en metodología para las posibilidades de curvas de amplificación y finalmente, poder calcular su proporción genotípica y frecuencia alélica para cada alelo (C o T).

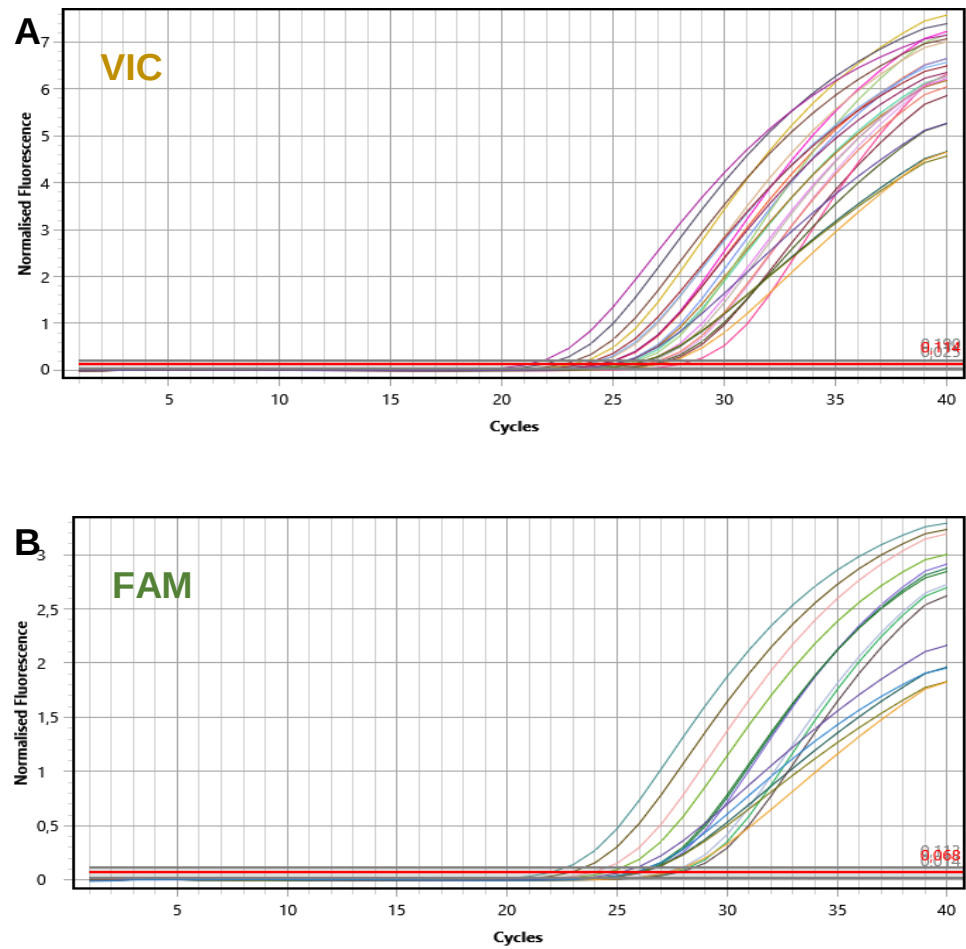


Figura 12. Curvas de amplificación para muestras rs2285666. Para A) Canal VIC. y B) Canal FAM.

Tabla 10. Genotipos obtenidos del total de muestras (n=98).

MUESTRA	GENOTIPO rs2285666 (ACE2)	GENOTIPO rs12329760 (TMPRSS2)	MUESTRA	GENOTIPO rs2285666 (ACE2)	GENOTIPO rs12329760 (TMPRSS2)	MUESTRA	GENOTIPO rs2285666 (ACE2)	GENOTIPO rs12329760 (TMPRSS2)	MUESTRA	GENOTIPO rs2285666 (ACE2)	GENOTIPO rs12329760 (TMPRSS2)
S8	TT	CC	S55	CC	CT	S108	CC	CC	P20	CC	CC
S10	CC	CC	S58	CC	CC	S110	TT	CC	P21	CC	CC
S12	CT	CC	S59	CC	CC	S111	CC	CC	P23	CC	CC
S13	CC	CC	S61	CC	CC	S116	CC	CC	P24	CC	CC
S15	CT	CT	S62	CC	CC	S119	TT	CC	P25	CT	CC
S21	CC	CT	S64	CC	CC	S120	CC	CC	P27	CT	CC
S22	CT	CT	S65	CC	CC	S121	CC	CC	P30	CC	CC
S26	CC	CT	S69	CC	CC	S122	CT	CT	P31	CC	CC
S28	CC	CC	S70	CT	CC	S124	CC	CC	P32	CC	CC
S30	CC	CT	S71	TT	CC	S129	TT	CC	P35	CC	CC
S32	CC	CC	S72	TT	CC	S130	TT	CC	P39	TT	CC
S37	TT	CC	S75	CC	CC	S131	CT	CC	P40	CC	CC
S40	CT	CC	S76	CC	CC	S133	CT	CC	P41	CC	CC
S41	CC	CC	S79	CT	CC	P2	CC	CC	P42	TT	CC
S42	TT	CT	S82	CC	CC	P4	CC	CC	P43	CC	CC
S43	CC	CC	S88	CC	CC	P5	CT	CC	P45	CC	CC
S44	TT	CC	S89	CC	CC	P7	CC	CT	P49	CT	CT
S45	CC	CC	S91	CC	CC	P12	CC	CC	P50	CT	CC
S46	TT	CC	S95	CC	CC	P13	CC	CC	P51	CC	CC
S47	CT	CC	S96	TT	CC	P14	CT	CT	P53	CC	CC
S48	CC	CC	S100	TT	CC	P15	TT	CC	P55	CC	CC
S49	CC	CC	S105	TT	CC	P16	CC	CC	N1	CC	CC
S53	CC	CC	S106	CC	CC	P17	CC	CC	N2	CC	CC
S54	CC	CC	S107	CC	CC	P19	TT	CC	N5	CT	CC
									A1	CC	CC
									A2	CT	CC

8.5. Determinación de distribución y frecuencia genotípica y alélica

Mediante la realización de PCR en tiempo real para genotipificación, se determinaron los polimorfismos en las 98 muestras de ADN de una población de la región del Biobío. Del total de la población, un 63,27% presentó el genotipo homocigoto *Wild Type* (CC), un 18,73% presentó el genotipo heterocigoto mutado (CT) y de igual manera, un 18,73% portó los dos alelos mutados (TT) para rs2285666 (Figura 13.A). Para rs12329760, un 88,78% presentó el genotipo homocigoto *Wild Type* (CC), un 11,22% presentó el genotipo heterocigoto mutado (CT) y no se encontró la portación de los dos alelos mutados (TT) en la población (Figura 13.B). Estos datos de distribución genotípica se muestran en la figura 13.

En cuanto a la distribución alélica, un 72,45% presentó el alelo *Wild Type* (C) y un 27,55% presentó el alelo de riesgo (T) para el polimorfismo rs2285666. Para rs12329760, un 94,39% presentó el alelo *Wild Type* (C) y un 5,61% presentó el alelo de riesgo (T) (Figura 14).

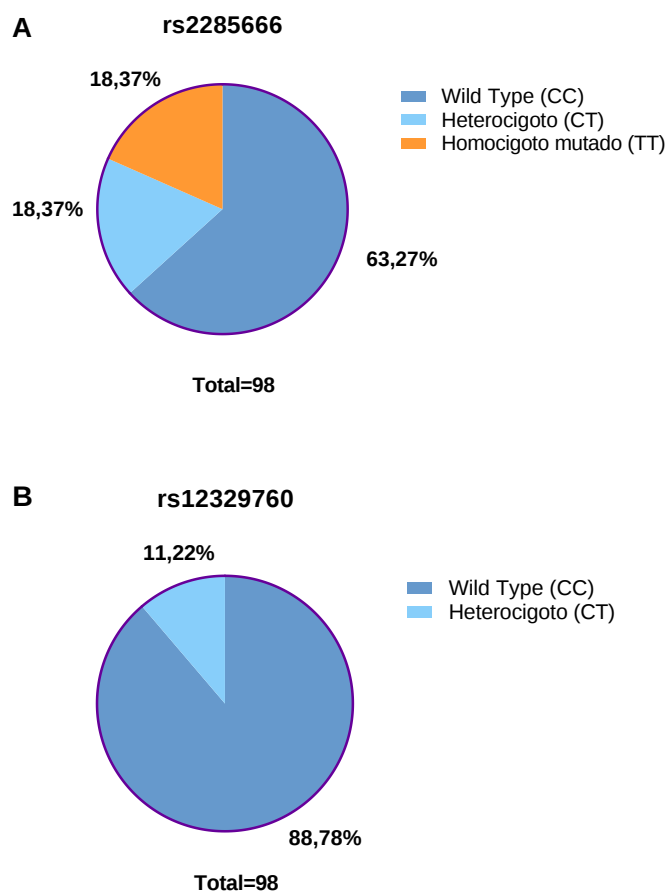


Figura 13. Distribución genotípica de las muestras de la población de estudio.

A) rs2285666 y B) rs12329760.

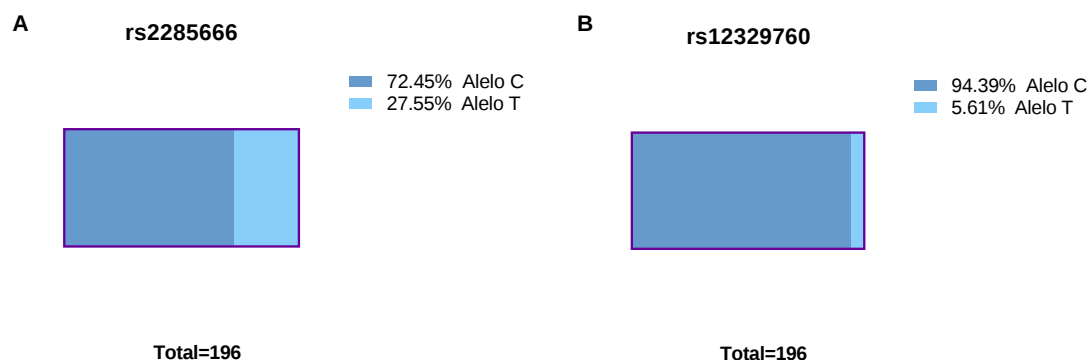


Figura 14. Distribución alélica de las muestras de la población de estudio.

A) rs2285666 y B) rs12329760.

Luego, se analizaron las frecuencias alélicas y genotípicas observadas y esperadas, para analizar el comportamiento de la población para ambos polimorfismos con el Equilibrio de Hardy Weinberg. Se observó para la distribución por genotipos para el polimorfismo rs2285666, una diferencia estadísticamente significativa ($p < 0,0001$) entre las frecuencias alélicas esperadas y observadas, indicando que no cumpliría con el Equilibrio de Hardy Weinberg. Para la población de rs12329760, no se observan diferencias significativas entre las frecuencias esperadas y observadas, ya que ambas son iguales. Por lo tanto, se concluye que esta última población, si cumple con el Equilibrio de Hardy Weinberg, como se observa en la tabla 11.

Tabla 11. Cálculo de frecuencias alélicas observadas y esperadas para ambos polimorfismos. Valores estadísticamente significativos: ns $p>0,05$; **** $p<0,0001$.

Genotipo	No	C	T	Frecuencias genotípicas observadas	Frecuencias genotípicas esperadas	Equilibrio Hardy Weinberg
rs2285666						
CC	62	124	0	0,64	0,52	No (p<0,0001)
CT	18	18	18	0,18	0,40	
TT	18	0	36	0,18	0,08	
Total	98	142	34	1	1	
Frecuencias alelicas		0,72	0,28			
rs12329760						
CC	87	174	0	0,88	0,88	Si (p=0,579)
CT	11	11	11	0,12	0,12	
TT	0	0	0	0	0	
Total	98	185	11	1	1	
Frecuencias alelicas		0,94	0,06			
	Total alelos	196				

8.6. Asociación entre variables clínicas, patologías y genotipos

En cuanto al análisis de los grupos de estudio, clasificados según sanos (n=48), obesidad (n=21), hipertensión (n=7), diabetes (n=5) y 2 o más patologías (n=17) y su análisis con los genotipos obtenidos, se observaron las siguientes distribuciones genotípicas, expresadas en cantidad de participantes que presentaron cada genotipo (CC, CT, TT), porcentaje (%) de un total, odds ratio (OR) y sus respectivos intervalos de confianza (CI 95%), además de su correspondiente *p* value para esta asociación. Lo anterior, también se realizó para los alelos (C y T). Cabe destacar que, para este análisis, se compararon las patologías con el grupo control (sanos) usando como genotipo y alelo de referencias, CC y C, respectivamente. No se encontraron diferencias significativas ($p > 0,05$) para ambos polimorfismos (rs2285666, tabla 12.A y rs12329760, tabla 12.B) entre los grupos de estudio y los genotipos. Además, se observa una tendencia en la distribución de los genotipos para la mayoría de los grupos (alto % CC, medio % CT, bajo % TT), a excepción del grupo de los hipertensos para el polimorfismo rs2285666 (tabla 12.A), en donde se observó un 42,9% que tiene el genotipo CT vs un 28,6% para CC y 28,5% TT, con un *p* value no significativo presenta una distribución diferente a los otros grupos.

Tabla 12. Genotipo y análisis estadístico de los genotipos vs patologías para

A) rs2285666 y B) rs12329760.

A

Variables	Control (Sanos)		Obesos		OR (CI 95%)	p value	Hipertensión		p value	Diabetes		OR (CI 95%)	p valu	2 o más patologías		OR (CI 95%)	p value		
	N°	%	N°	%			N°	%		N°	%			N°	%				
Genotipo ACE2 (rs2285666)																			
CC	31	64,6	15	71,4	1,0 (Ref)	-	2	28,6		-	2	40	1,0 (Ref)	-	12	70,6	1,0 (Ref)	-	
CT	7	14,6	4	19,0	1,18 (0,34 - 4,91)	0,99	3	42,9	6,64 (1,11 - 40,38)	0,07	1	20	2,21 (0,14 - 20,67)	0,48	3	17,6	1,11 (0,28 - 4,38)	0,99	
TT	10	20,8	2	9,6	0,41 (0,08 - 1,99)	0,48	2	28,5	3,10 (0,43 - 21,12)	0,27	2	40	3,10 (0,43 - 21,12)	0,29	2	11,8	0,52 (0,10 - 2,63)	0,70	
Alelos ACE2 (rs2285666)																			
C	69	71,9	34	81,0	1,0 (Ref)	-	7	50,0		1,0 (Ref)	-	5	50	1,0 (Ref)	-	27	79,4	1,0 (Ref)	-
T	27	28,1	8	19,0	0,60 (0,26 - 1,50)	0,29	7	50,0	2,56 (0,87 - 7,42)	0,12	5	50	2,55 (0,74 - 8,72)	0,16	7	20,6	0,66 (0,26 - 1,66)	0,50	

B

Variables	Control (Sanos)		Obesos		OR (CI 95%)	p value	Hipertensión		OR (CI 95%)	p value	Diabetes		OR (CI 95%)	p value	2 o más patologías		OR (CI 95%)	p value
	N°	%	N°	%			N°	%			N°	%						
Genotipo TMPRSS2 (rs12329760)																		
CC	42	87,5	18	85,7	1,0 (Ref)	-	5	71,4	1,0 (Ref)	-	5	100	1,0 (Ref)	-	17	100	1,0 (Ref)	-
CT	6	12,5	3	14,3	1,17 (0,29 - 4,41)	0,99	2	28,6	2,80 (0,46 - 14,17)	0,27	0	0	-	-	0	0	-	-
TT	0	0	0	0	-	-	0	0	-	-	0	0	-	-	0	0	-	-
Alelos TMPRSS2 (rs12329760)																		
C	90	96,1	39	92,9	1,0 (Ref)	-	12	85,7	1,0 (Ref)	-	10	100	1,0 (Ref)	-	34	100	1,0 (Ref)	-
T	6	3,9	3	7,1	1,15 (0,30 - 4,31)	0,99	2	14,3	2,50 (0,47 - 12,07)	0,27	0	0	-	-	0	0	-	-

Posteriormente, se realizó un análisis de componentes principales (PCA) con el fin de representar la distribución de las muestras dentro del espacio de los componentes principales, de acuerdo a los datos utilizados. Se utilizó el método UMAP, con reducción de dimensionalidad (Figura 15.A y 15.B). No se observó agrupación espontánea y no existe tendencia en la agrupación por parte de las muestras al evaluar sus clases para los dos polimorfismos (rs2285666 clase 0: CC, clase 1: CT y clase 2: TT). Lo anterior es coherente con el análisis univariado que se realizó analizando el genotipo y el alelo de referencia del grupo control, no encontrando diferencias significativas entre ambos grupos (no existen características de que tiendan a agrupar a los genotipos y que sea una clasificación relevante).

A pesar de no encontrar clusters de diferenciación, se procedió a buscar la importancia de cada característica para el modelo de predicción de separación de las clases (genotipos). Esto proporciona una visión general de las características más influyentes para el modelo y su impacto en la clasificación de cada clase. Se utilizó un modelo de clasificación usando herramientas de machine learning. Se balancearon las clases por la técnica SMOTE y se utilizó el método SHAP. En el eje X se ubican las características de mayor a menor importancia. En el eje Y se grafica el promedio del valor absoluto del valor de SHAP, los colores identifican las clases (genotipos) y la longitud de las barras

indican la influencia de las características en la predicción del modelo. Para el polimorfismo rs2285666 (figura 16.A), de todas las variables, la talla es la característica más importante para poder encontrar una diferencia entre las clases (genotipos), seguido del IMC y el género. Para el polimorfismo rs12329760, las características más importantes son el género, la edad y el IMC. (figura 16.B).

Finalmente, se realizó un gráfico de resumen que combina la importancia de las características con los efectos de la característica. Cada punto en la figura X.A y X.B resumen es un valor de Shapley para una entidad y una instancia. La posición en el eje Y está determinada por la característica y en el eje X por el valor de Shapley. El color representa el valor de la característica mayor a menor. Los puntos superpuestos fluctúan en la dirección del eje Y, por lo se tiene una idea de la distribución de los valores Shapley por entidad. Las características se ordenan según su importancia. (Figura 17.A y 17.B). De este gráfico se obtuvo que para el polimorfismo rs2285666 (Figura 17.A), las tres primeras características de mayor importancia (IMC, talla y edad), poseen valores de Shapley tanto altos como bajos en ambos lugares de la gráfica (valores negativos y positivos), lo que podría indicar que, para estas características, no hay valores (ya sea altos o bajos) que contribuyan al modelo de predicción de separación de genotipos. Se puede observar una leve agrupación de valores bajos para la edad

(puntos azules), que podría relacionar el aumento de la edad con el impacto negativo en la separación de las clases. Se observa una agrupación de valores bajos para las patologías de hipertensión arterial y diabetes, presentando así una contribución negativa al modelo de predicción para los valores bajos de estas características. Lo mismo anterior ocurre para el polimorfismo rs12329760, con la diferencia que la característica que más importa para predecir el modelo (generar las clases) es el género, al igual como se representa en la figura 16.B.

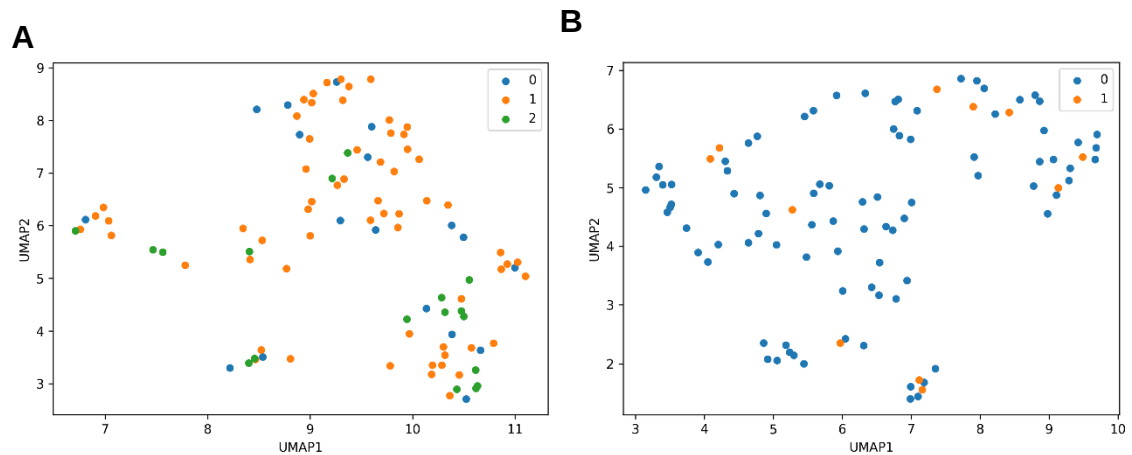


Figura 15. Análisis de componentes principales (PCA) para los polimorfismos.

A) rs2285666 y B) rs12329760.

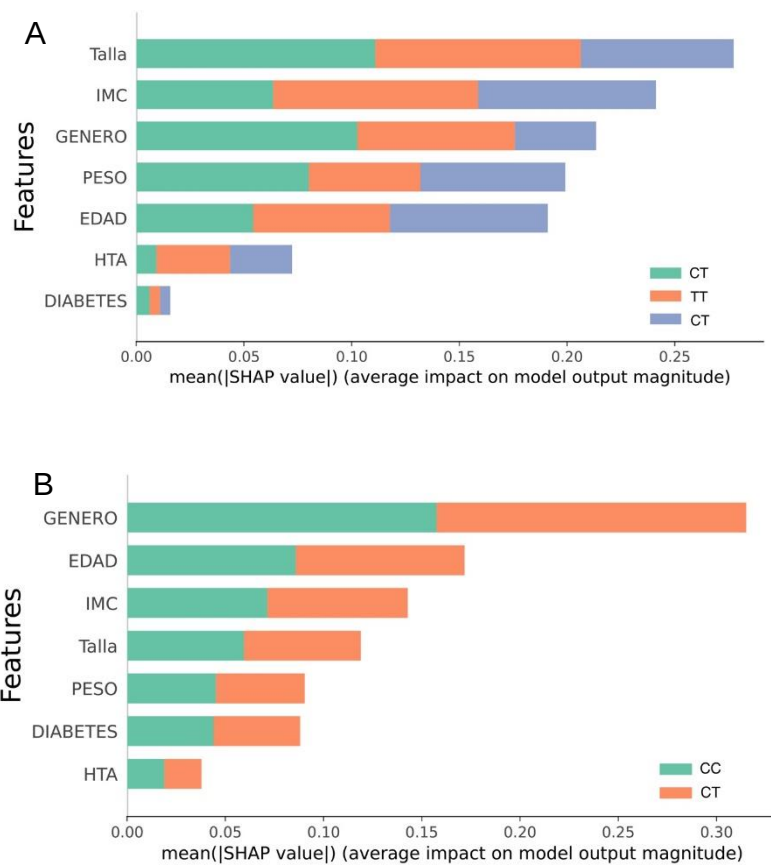


Figura 16. Análisis de método de SHAP para los polimorfismos. A) rs2285666 y B) rs12329760.

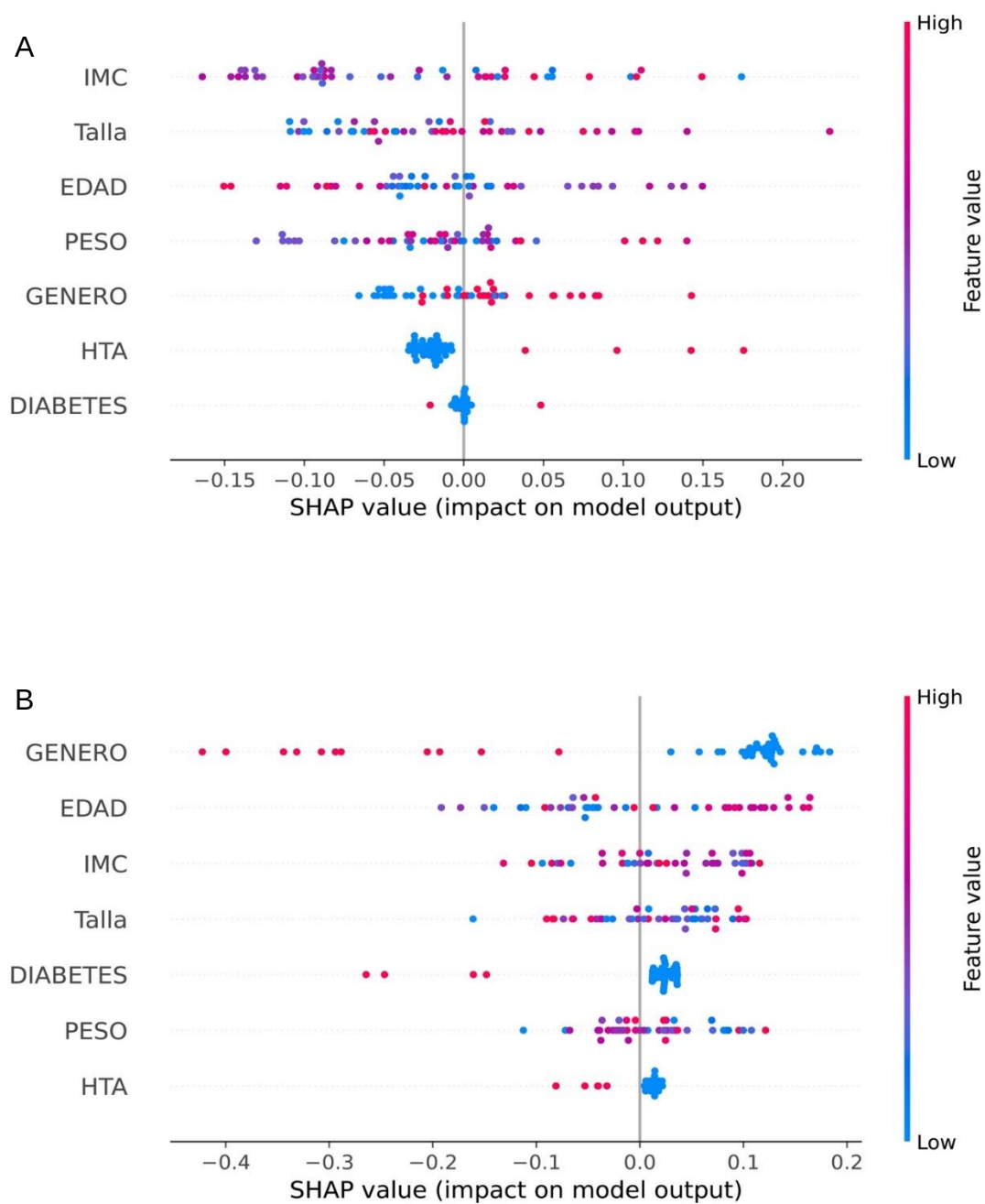


Figura 17. Beeswarm plot para los polimorfismos. A) rs2285666 y B) rs12329760.

9. DISCUSIÓN

El presente trabajo de investigación en Internado en Sector Salud, permitió establecer la presencia de genotipos y alelos de riesgo para dos polimorfismos en población chilena: rs2285666 y rs12329760, ambos asociados a gravedad por COVID-19. Este es el primer estudio realizado en población chilena de alelos de riesgos asociados a la infección por SARS-CoV-2.

Los polimorfismos se han asociado ampliamente a diversas enfermedades. Los investigadores han utilizado varios enfoques para identificar la variación genética que predispone a las personas a enfermedades comunes. Un enfoque más sensible para identificar alelos de riesgo con efectos genéticos más pequeños es emplear un estudio de asociación de casos y controles en el que se usan marcadores alélicos, como los SNP, para encontrar regiones del genoma enriquecidas (o empobrecidas) en un alelo o haplotipo de riesgo particular. entre los casos y los controles. Si ciertas variaciones genéticas son más frecuentes en personas con la enfermedad, se dice que estas variaciones están "asociadas" con esa enfermedad. Estas variaciones son entonces consideradas como señalizadores de la región del genoma humano donde probablemente esté el problema causante de la enfermedad. Siguiendo esta línea, la mayoría de las

variaciones de SNPs asociadas con enfermedades no se presentan en las regiones codificantes del ADN, sino que normalmente se sitúan en las amplias regiones no codificantes intergénicas o en los intrones que se extraen de la secuencia de ADN cuando las proteínas son procesadas. Estas son presumiblemente secuencias de ADN que controlan otros genes, pero normalmente su función proteica no se conoce [45]

Lo anterior, fue la motivación para generar este estudio, ya que desde que comenzó la pandemia por COVID-19, existían muchas interrogantes acerca de los factores que podían contribuir a la gravedad de la enfermedad, muchos de ellos poco conocidos incluyendo los factores genéticos. El primer paso para intentar responder a esta interrogante, fue realizar una revisión sistemática de los polimorfismos que se encontrara más evidencia científica al estar relacionada con la enfermedad. Desde el año 2020, comenzaron a aparecer los primeros estudios *in silico*, que trataron de explicar mediante softwares, la estimación de los cambios en la diferencia de energía libre ($\Delta\Delta G$) lo que podría indicar una mayor afinidad de unión a la proteína S y, por lo tanto, una mayor susceptibilidad a infección por SARS-CoV-2 de algunas variantes de ACE2 [46]. También, se realizaron estudios de modelamiento molecular (Molecular Docking), estimando el *Global Energy Score* (Kcal/mol) de la unión de las diferentes variantes con la proteína S de SARS-CoV-2 [47]. Todo lo anterior se realizó en conjunto con

estudios poblacionales de frecuencias alélicas (MAF: *minor allele frequencies*) para los países que se veían más afectados por la pandemia en ese momento.

Posteriormente para el año 2021, se comenzaron a realizar los primeros estudios *in vitro* de estas variantes identificadas anteriormente mediante softwares computacionales. Los estudios comprendían análisis de ensayos de afinidad, ELISA y de colorimetría, confirmando lo que se predijo anteriormente para las variables de ACE2 con los estudios *in silico* [32]. En esos momentos se comenzaba a conocer la fisiopatología de la entrada del virus SARS-CoV-2 a células humanas, identificándose la enzima convertidora de angiotensina 2 como receptor. Por lo anterior, los estudios fueron dirigidos hacia esa enzima, la cual ha sido ampliamente estudiada e involucrada a polimorfismos asociados a hipertensión [48]. Por otro lado, se conoció que la proteasa serina transmembrana 2 (TMPRSS2) tenía un rol fundamental en la entrada del virus a la célula. Esta proteasa, la cual no se conoce realmente su función, ha sido vinculada a ciertos polimorfismos con riesgo de cáncer de próstata [49].

Posteriormente para el año 2022, surgieron los primeros estudios *in vivo* que venían a apoyar o a refutar lo que se estaba comenzando a describir de estas dos proteínas claves en la entrada del virus. Éstas se han investigado en varios estudios, aunque sus resultados son controvertidos. Algunos estudios han

correlacionado el alelo T (A) de los SNP ACE2 rs2285666 [50] y TMPRSS2 rs12329760 [51] con una menor gravedad de COVID-19 en una población india. Estos datos también han sido replicados en un estudio multinacional [52] y confirmados por el metanálisis de *Dobrijevic et al., 2022* concluyendo que los alelos T de ACE2 rs2285666 (OR 0.512, 95% IC 0.331–0.793) y TMPRSS2 rs12329760 (OR 0.734, 95% IC 0.560–0.960) SNP se asociaron con un menor riesgo de desarrollar COVID-19 grave [53]. Otro metanálisis encontró que TMPRSS2 rs12329760 no se asoció con manifestaciones más graves de COVID-19, mientras que los SNP ACE2 rs2285666 y ACE2 rs2074192 se asociaron con manifestaciones más graves de COVID-19 por genotipo, pero no por alelo [54].

Por otro lado, otros estudios han observado resultados diferentes. *Martínez-Gómez et al., 2022* han identificado recientemente que el alelo T de ACE2 rs2285666 se asoció con un mayor riesgo de resultados críticos (OR 1.89, 95% IC 1.06–3.35) de COVID-19, especialmente para los hombres [55]. *Wulandari et al., 2021* observó que el TMPRSS2 rs12329760 estaba asociado con una mayor carga viral pero no con una condición aguda más grave [56]. El metanálisis de *Saengsiwaritt et al., 2022*. reveló una asociación entre el alelo C de TMPRSS2 rs12329760 (OR 1,32, IC del 95 % 1,01–1,73) y el alelo T de ACE2 rs2285666 (OR 2,14, IC del 95 % 1,26–3,66) con un mayor riesgo de COVID-19 grave [57]. La evidencia actual sugiere una clara influencia de estos polimorfismos en la

COVID-19, pero hace más compleja la subestimación de su asociación con la fisiopatología de la infección. Se podría entender de todo lo reportado anteriormente que estas variantes podrían ser funcionales, en donde podría estar aumentada la expresión génica y esto asociarse con una mayor cantidad de sitios de unión viral unidos a la membrana, lo que aumenta la vulnerabilidad de los portadores a la infección. Estas variantes de riesgo pueden ser particularmente adversas en los hombres, que portan solo una copia del gen ACE2 ligado al cromosoma X. Cabe recordar que el SNP rs2285666 está ubicado en los nucleótidos de empalme intrónicos y, por lo tanto, podría afectar el procesamiento del ARN total de ACE2 a ARNm y, finalmente, la cantidad de proteína. Al menos un estudio ha informado niveles más altos de ACE2 circulante en hombres que en mujeres [58]. Un SNP intrónico diferente en fuerte desequilibrio de enlace con rs2285666 (rs879922) se ha asociado con la expresión del gen ACE2 [59]. Para TMPRSS2 (rs12329760), el escenario es un poco distinto.

El polimorfismo rs12329760 se han correlacionado con la susceptibilidad y la gravedad de la infección por el virus de la influenza H1N1 o H1N9 [60]. Por el momento, existen pocas revisiones de la literatura e investigaciones *in silico* basadas en datos de genotipado de bases de datos existentes [61, 62]. Debido a su importancia funcional, la variante sin sentido rs12329760, que cambia el aminoácido valina en la posición 160 a metionina, podría ser importante en el

contexto de COVID-19 [63]. El alelo T rs12329760 está relacionado con las numerosas copias del gen de fusión TMPRSS2-ERG, que está relacionado con una peor supervivencia y una mayor recurrencia tumoral en el cáncer de próstata [64]. Existen pocos estudios *in vivo* para este polimorfismo, sin embargo, la mayoría asocia el alelo de riesgo T con enfermedad grave por COVID-19. Yaghoobi A, et al., 2023. investigó la asociación entre el genotipo TMPRSS2 y la gravedad de la enfermedad por coronavirus 2019 (COVID-19) en pacientes iraníes. Los resultados mostraron una asociación significativa entre el alelo T menor y la gravedad del COVID-19 (valor $p = 0,043$) bajo el modelo de herencia dominante y aditivo. En conclusión, los resultados de este estudio mostraron que el alelo T del rs12329760 en el gen TMPRSS2 es un alelo de riesgo para la forma grave de COVID-19 en pacientes iraníes. Sin embargo, ellos postulan que se necesitan más estudios para abordar los complejos mecanismos detrás de la interacción de la proteína TMPRSS2 y el SARS-CoV-2 y el papel del polimorfismo rs12329760 en la determinación de la gravedad de la enfermedad [65].

En base a los resultados de esta investigación, los cuales se limitaron a conocer si estaba presente o no este polimorfismo, calcular las distribuciones genotípicas y frecuencias alélicas, mostraron que para nuestra población total ($n=98$), los resultados indicaron distribuciones genotípicas de 63,27% CC, 18,73% CT y 18,73% TT para rs2285666. Un 88,78% CC, 11,22% CT y 0% TT para

rs12329760. En relación a la frecuencia alélica, se encontraron frecuencias relacionadas al alelo de riesgo (T) de 0,28 para rs2285666 y 0,06 para rs12329760. Estos resultados no coinciden con los reportados en el algoritmo 1000 Genomes Project para población americana (C= 0,66 y T= 0,34, para rs2285666 y C=0,85 y T=0,15).

En relación con el estudio del equilibrio de Hardy Weinberg, que está basado en las frecuencias alélicas y genotípicas de una población donde este se define en que estas frecuencias se mantengan constantes de generación en generación, es decir, se encuentren en equilibrio a menos que se introduzcan influencias perturbadoras [66]. Este equilibrio genético es un estado ideal que ofrece una línea para medir el cambio entre generaciones, que es imposible en la naturaleza. Con la ley de Hardy-Weinberg se asentaron los cimientos de la genética de poblaciones, según la cual, la alteración genética de una población sólo puede darse por factores como mutaciones, selección natural, influencias casuales, convergencias o divergencias individuales, de modo que el cambio genético implica la perturbación del equilibrio establecido por la ley de Hardy-Weinberg. Se observó que para rs2285666 las frecuencias alélicas observadas son significativamente distintas de las esperadas según este equilibrio, incumpliendo con el equilibrio. Esto se puede visualizar concretamente en la misma cantidad que se obtuvo del genotipo heterocigoto mutado (CT) y del homocigoto mutado

(TT), lo cual no es esperado para este equilibrio. Para el polimorfismo rs12329760, se pudo comprobar el equilibrio de Hardy Weinberg.

Para poder explicar bilateralmente ambos casos, se sabe que para este equilibrio deben cumplirse ciertos supuestos: ausencia de mutación y migración, gran tamaño de la población, apareamiento al azar, y ausencia de selección natural. Con lo descrito anteriormente, se conoce al gen de ACE2 con una variedad de polimorfismos, lo que no es el caso del gen de TMPRSS2. Uno de los supuestos que podríamos hipotetizar que no se estaría cumpliendo para el gen de ACE2 es el concepto de mutación dado por un término llamado deriva genética, el cual es un mecanismo de evolución caracterizado por fluctuaciones aleatorias en la frecuencia de una versión determinada de un gen (alelo) en una población. Los efectos de la deriva genética pueden ser fuertes y a veces llevan a que determinados rasgos se vuelvan abrumadoramente frecuentes o a que desaparezcan de una población. Otro supuesto del equilibrio es el tamaño de la población, ya que como se dijo anteriormente, un $n=98$ es un número bastante pequeño para estudio de genética de poblaciones, además, los grupos de estudio poseían grandes diferencias en número. Sin embargo, otro factor importante a considerar en el equilibrio de Hardy Weinberg, es la aleatoriedad de las muestras por lo que al realizar el siguiente paso de esta investigación (convocar a grupos seleccionados con características en comunes, patologías) es probable que el

equilibrio tampoco se cumpla. No obstante, el cumplimiento de esta ley no genera cambios en el análisis, solo establece que la población estudiada se mantendrá constante de generación en generación con la ausencia de factores perturbadores.

En relación al análisis de los genotipos con las patologías, no se encontraron asociaciones entre los genotipos de ambos polimorfismos con los grupos de estudio. Es decir, no existe una relación como la que se ha reportado para polimorfismos de ACE2 para hipertensión, con los analizados en este estudio. Esto puede tener relación con el desarrollo de la enfermedad o con su fisiopatología, ya que los polimorfismos estudiados no están involucrados directamente en los procesos patológicos. Como por ejemplo, los estudios de obesidad y genética se han centrado en genes relacionados con el control del hambre y almacenaje de grasa en las células, como lo son el LEP, el LEPR y el MC4R del sistema leptina-melanocortina, el cual opera en la regulación neuroendocrina de la ingestión y el equilibrio energético e influye en la patogenia de la enfermedad [67]. Para la diabetes, se han publicado numerosos informes sobre la genética de la DM2, y los más recientes muestran el efecto de los SNP en varios genes correspondientes a la predicción del riesgo de DM2, como las variantes genéticas del receptor gamma activado por el proliferador de peroxisomas (PPAR- γ), [68] enzima convertidora de angiotensina (ACE) [69],

metileno tetrahidrofolato reductasa (MTHFR) [70], proteína de unión a ácidos grasos 2 (FABP2) [71] y gen asociado a la masa grasa y la obesidad (FTO) [72]. Para la hipertensión, se han asociado polimorfismos a ACE2 y a los receptores de angiotensina 2 tipo 1 y tipo 2 [73]. Sin embargo, de los polimorfismos asociados a hipertensión no está reportado rs2285666.

Finalmente, es necesario recordar que el conocimiento de nuevos marcadores biológicos (genéticos, en este caso), es necesario un amplio trabajo y análisis de muchos factores que pudiesen contribuir en el desarrollo o progresión de una enfermedad. Cabe recordar, que la OMS declaró fin de la pandemia el día 5 de mayo de 2023, tras más de 3 años del comienzo de esta. Durante este tiempo, la comunidad científica ha realizado esfuerzos enormes por entender como el virus del SARS-CoV-2 tuvo tanto impacto a nivel mundial y produjo millones de muertes, sin embargo, existen interrogantes que aún no se responden o que falta tiempo para que se respondan. Una de esas interrogantes es como los factores genéticos contribuyen a esta enfermedad. Es probable que los genes posean muchos polimorfismos y que los estudios apunten a ciertos en particular, pero se debe evaluar la evolución de las poblaciones y que están en constante mutación, la que puede generar aparición de nuevos polimorfismos que desencadenen otro tipo de acción en la fisiopatología de cualquier enfermedad.

10. CONCLUSIÓN Y PROYECCIONES

Se logró determinar la presencia del polimorfismo rs2285666 y rs12329760 en una población de la región del Bío-bío, presentando genotipos y alelos de riesgo para rs2285666 y alelos de riesgo para rs12329760. De igual manera, se logró determinar que la población en estudio, clasificada en grupos controles (sanos) y que presentan alguna patología (obesidad, diabetes, hipertensión y 2 o más de estas patologías), no se relacionan significativamente con alguno de estos polimorfismos. Cabe destacar que la búsqueda de esta asociación, tiene que ver con el sin número de patologías que son relacionadas con diversos polimorfismos. Sin embargo, a pesar de no obtener diferencias significativas, se pudo observar una tendencia en el genotipo CT para el grupo de los participantes que reportaron hipertensión y que indica que podría existir una relación entre el genotipo rs2285666 (ACE2) y su desarrollo. Lo anterior se condice con el rol fundamental que posee ACE2 en el sistema renina-angiotensina-aldosterona para el control de la presión arterial y con la asociación de una gran cantidad de polimorfismos asociados al desarrollo de esta patología.

Como limitaciones de este estudio, se puede observar que el número de participantes (n=98) fue reducido para realizar un estudio poblacional de

polimorfismos, los cuales sobrepasan los $n=200$ en la mayoría de los casos. También, se deriva de esto, la poca cantidad de voluntarios que entraban dentro de los grupos con obesidad, hipertensión y/o diabetes. Lo anterior podemos explicarlo con el Anexo 4 (formulario de autoreporte de datos de que respondieron los voluntarios), subestimando patologías que, al momento de responder la encuesta, no tenían conocimiento de su estado de salud.

Finalmente, con lo reportado anteriormente, es fundamental poder hacer cambios en la metodología de este estudio para dirigir el análisis. En primera instancia, poder reclutar una mayor cantidad de voluntarios, aumentando el número de participantes que presenten patologías y que así este grupo sea más representativo. En este caso, sería ideal reclutar voluntarios que ya tengan una patología diagnosticada, aunque de igual forma, poder medir parámetros (glicemia, presión arterial, cálculo de IMC) y así obtener datos más concretos que no se puedan estar subestimando como el reporte de una patología. En concreto, luego de conocer que existen alelos de riesgo en población chilena para enfermar gravemente por COVID-19, se sugiere dirigir la población de estudio a la identificación de estos polimorfismos en pacientes sanos, moderados y que hayan enfermado gravemente, para así poder correlacionar estas variables y conocer si los alelos de riesgo de estos polimorfismos podrían ser predictores de gravedad por COVID-19 en Chile.

11. GLOSARIO

ACE2: *Angiotensin converting enzyme 2*

ADN: Ácido desoxirribonucleico

ACE: *Angiotensin-converting enzyme*

ACT2: *Alveolar cells type 2*

Ang (1-7): Angiotensina 1-7

Ang (1-9): Angiotensina 1-9

Ang I: Angiotensina I

Ang II: Angiotensina II

ARN: Ácido ribonucleico

ARNm: Ácido ribonucleico mensajero

B^oAT1: *Sodium-dependent neutral amino acid transporter*

CESFAM: Centro de Salud Familiar

CLD: *Collectrin domain*

COVID-19: *Coronavirus Disease 19*

CoVs: Coronavirus

EDTA: Ácido etilendiaminotetraacético

ERG: *Erythroblast transformation-specific gene*

ETS: *Erythroblast transformation-specific*

ETV1: *ETS variant transcription factor 1*

EU: Enfermera Universitaria

HEXXH: Dominio de unión a zinc

HNF: Hisopado nasofaríngeo

HTA: Hipertensión arterial

IMC: Índice de masa corporal

ISP: Instituto de Salud Pública

MAF: *Minor allele frequency*

MATPEL: Empresa materiales peligrosos

MBG: *Minor groove binding*

MERS-CoV: *Middle East respiratory syndrome-Coronavirus*

Nal: Yoduro de Sodio

OR: *Odds ratio*

PCA: *Principal components analysis*

PCR: *Polymerase chain reaction*. Proteína C reactiva

PD: *Peptidase domain*

qPCR: *Real-time polymerase chain reaction*

RAAS: *Renin-angiotensin-aldosterone system*

RBD: *Receptor binding domain*

RIC: Rango intercuartílico

SARS-COV: *Severe acute respiratory syndrome coronavirus 2*

SARS-COV-2: *Severe acute respiratory syndrome coronavirus*

SD: *Standard deviation*

SHAP: *Shapley additive exPlanations*

SMOTE: *Synthetic minority oversampling technique*

SNP: *Single nucleotid polymorphism*

TAE: Tris-acetato-EDTA

TE: Tris-EDTA

TMPRSS2: *Transmembrane protease serine 2*

UMAP: *Uniform manifold approximation and project*

UV: Ultravioleta

12. BIBLIOGRAFÍA

1. Zhu, N., et al., *A Novel Coronavirus from Patients with Pneumonia in China, 2019*. N Engl J Med, 2020. **382**(8): p. 727-733.
2. Li, F., *Structure, Function, and Evolution of Coronavirus Spike Proteins*. Annu Rev Virol, 2016. **3**(1): p. 237-261.
3. Zaki, A.M., et al., *Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia*. N Engl J Med, 2012. **367**(19): p. 1814-20.
4. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)". ArcGIS. Johns Hopkins University. Recuperado 22 de mayo 2023.
5. Adhikari, S.P., et al., *Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review*. Infect Dis Poverty, 2020. **9**(1): p. 29.
6. Rothe, C., et al., *Transmission of 2019-nCoV Infection from an Asymptomatic Contact in Germany*. N Engl J Med, 2020. **382**(10): p. 970-971.
7. Lu, R., et al., *Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding*. Lancet, 2020. **395**(10224): p. 565-574.

8. Wan, Y., et al., *Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus*. J Virol, 2020. **94**(7).
9. Zou, X., et al., *Single-cell RNA-seq data analysis on the receptor ACE2 expression reveals the potential risk of different human organs vulnerable to 2019-nCoV infection*. Front Med, 2020. **14**(2): p. 185-192.
10. Zhang, H., Kang, Z., Gong, H., Xu, D., Wang, J., Li, Z., ... & Xu, H. (2020). The digestive system is a potential route of 2019-nCoV infection: a bioinformatics analysis based on single-cell transcriptomes. BioRxiv.
11. Chai, X., Hu, L., Zhang, Y., Han, W., Lu, Z., Ke, A., ... & Lan, F. (2020). Specific ACE2 expression in cholangiocytes may cause liver damage after 2019-nCoV infection. BioRxiv.
12. Mousavizadeh, L. and S. Ghasemi, *Genotype and phenotype of COVID-19: Their roles in pathogenesis*. J Microbiol Immunol Infect, 2021. **54**(2): p. 159-163.
13. Chen, Y., et al., *Structure analysis of the receptor binding of 2019-nCoV*. Biochem Biophys Res Commun, 2020. **525**(1): p. 135-40.
14. Li, M.Y., et al., *Expression of the SARS-CoV-2 cell receptor gene ACE2 in a wide variety of human tissues*. Infect Dis Poverty, 2020. **9**(1): p. 45.

15. Rabi, F.A., et al., *SARS-CoV-2 and Coronavirus Disease 2019: What We Know So Far*. Pathogens, 2020. **9**(3).
16. Rabaan, A.A., et al., *SARS-CoV-2, SARS-CoV, and MERS-COV: A comparative overview*. Infez Med, 2020. **28**(2): p. 174-184.
17. Chaudhry, F., et al., *Manipulation of ACE2 expression in COVID-19*. Open Heart, 2020. **7**(2).
18. Harmer, D., et al., *Quantitative mRNA expression profiling of ACE 2, a novel homologue of angiotensin converting enzyme*. FEBS Lett, 2002. **532**(1-2): p. 107-10.
19. Li, Y., et al., *Physiological and pathological regulation of ACE2, the SARS-CoV-2 receptor*. Pharmacol Res, 2020. **157**: p. 104833.
20. Verma, A., et al., *Expression of Human ACE2 in Lactobacillus and Beneficial Effects in Diabetic Retinopathy in Mice*. Mol Ther Methods Clin Dev, 2019. **14**: p. 161-170.
21. Wong, S.H., R.N. Lui, and J.J. Sung, *Covid-19 and the digestive system*. J Gastroenterol Hepatol, 2020. **35**(5): p. 744-748.
22. National Center for Biotechnology Information. ACE2 angiotensin I converting enzyme 2 [Homo sapiens (human)] [Internet]. Citada 2021 octubre 25 <https://www.ncbi.nlm.nih.gov/gene/59272>.

23. Tipnis, S.R., et al., *A human homolog of angiotensin-converting enzyme. Cloning and functional expression as a captopril-insensitive carboxypeptidase*. J Biol Chem, 2000. **275**(43): p. 33238-43.
24. Yan, R., et al., *Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2*. Science, 2020. **367**(6485): p. 1444-1448.
25. Sharma, R.K., et al., *ACE2 (Angiotensin-Converting Enzyme 2) in Cardiopulmonary Diseases: Ramifications for the Control of SARS-CoV-2*. Hypertension, 2020. **76**(3): p. 651-661.
26. National Human Genome Research Institute. Polymorphism. [Internet]. Citada 2022 enero 25. <https://www.genome.gov/es/genetics-glossary/Polimorfismo>
27. Yi, L., et al., *Association of ACE, ACE2 and UTS2 polymorphisms with essential hypertension in Han and Dongxiang populations from north-western China*. J Int Med Res, 2006. **34**(3): p. 272-83.
28. Malard, L., et al., *The association between the Angiotensin-Converting Enzyme-2 gene and blood pressure in a cohort study of adolescents*. BMC Med Genet, 2013. **14**: p. 117.
29. Patnaik, M., et al., *Association of angiotensin-converting enzyme and angiotensin-converting enzyme-2 gene polymorphisms with essential*

hypertension in the population of Odisha, India. Ann Hum Biol, 2014. 41(2): p. 145-52.

30. Duru, K., et al., *Frequency of a deletion polymorphism in the gene for angiotensin converting enzyme is increased in African-Americans with hypertension. Am J Hypertens, 1994. 7(8): p. 759-62.*

31. Devaux, C.A., J.M. Rolain, and D. Raoult, *ACE2 receptor polymorphism: Susceptibility to SARS-CoV-2, hypertension, multi-organ failure, and COVID-19 disease outcome. J Microbiol Immunol Infect, 2020. 53(3): p. 425-435.*

32. Suryamohan, K., et al., *Human ACE2 receptor polymorphisms and altered susceptibility to SARS-CoV-2. Communications Biology, 2021. 4(1): p. 475.*

33. Stopsack, K.H., et al., *TMPRSS2 and COVID-19: Serendipity or Opportunity for Intervention? Cancer Discov, 2020. 10(6): p. 779-782.*

34. Rajput, A.B., et al., *Frequency of the TMPRSS2:ERG gene fusion is increased in moderate to poorly differentiated prostate cancers. J Clin Pathol, 2007. 60(11): p. 1238-43.*

35. Abdelsattar, S., et al., *ACE2 and TMPRSS2 SNPs as Determinants of Susceptibility to, and Severity of, a COVID-19 Infection. Br J Biomed Sci, 2022. 79: p. 10238.*

36. Möhlendick, B., Schönfelder, K., Breuckmann, K., Elsner, C., Babel, N., Balfanz, P., ... & Kribben, A. (2021). ACE2 polymorphism and susceptibility for

SARS-CoV-2 infection and severity of COVID-19. *Pharmacogenetics and genomics*, 31(8), 165-171.

37. Yang, M., Zhao, J., Xing, L., & Shi, L. (2015). The association between angiotensin-converting enzyme 2 polymorphisms and essential hypertension risk: A meta-analysis involving 14,122 patients. *Journal of the Renin-Angiotensin-Aldosterone System*, 16(4), 1240-1244.

38. Paniri, A., Hosseini, M. M., & Akhavan-Niaki, H. (2021). First comprehensive computational analysis of functional consequences of TMPRSS2 SNPs in susceptibility to SARS-CoV-2 among different populations. *Journal of Biomolecular Structure and Dynamics*, 39(10), 3576-3593.

39. Dadras, O., Afsahi, A. M., Pashaei, Z., Mojdeganlou, H., Karimi, A., Habibi, P., ... & SeyedAlinaghi, S. (2022). The relationship between COVID-19 viral load and disease severity: A systematic review. *Immunity, inflammation and disease*, 10(3), e580.

40. Wulandari, L., Hamidah, B., Pakpahan, C., Damayanti, N. S., Kurniati, N. D., Adiatmaja, C. O., ... & Oceandy, D. (2021). Initial study on TMPRSS2 p. Val160Met genetic variant in COVID-19 patients. *Human genomics*, 15(1), 1-9.

41. Cui, J., F. Li, and Z.L. Shi, *Origin and evolution of pathogenic coronaviruses*. *Nat Rev Microbiol*, 2019. **17**(3): p. 181-192.

42. Liu, Z., et al., *Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2*. J Med Virol, 2020. **92**(6): p. 595-601.
43. Salazar, L.A., et al., *Micrométodo para extração de DNA genômico útil no diagnóstico molecular da hipercolesterolemia familiar*. Rev bras anal clin, 2001. **33**(3): p. 111-6.
44. Yan, X., et al., *PL - 030 The effects of ACE gene polymorphisms on ACE content before and after High-Intensity Interval Exercise*. Exercise Biochemistry Review, 2018. **1**.
45. Pearson, T. A., & Manolio, T. A. (2008). How to interpret a genome-wide association study. *Jama*, 299(11), 1335-1344.
46. Ali, F., Elserafy, M., Alkordi, M. H., & Amin, M. (2020). ACE2 coding variants in different populations and their potential impact on SARS-CoV-2 binding affinity. *Biochemistry and biophysics reports*, 24, 100798.
47. Calcagnile, M., Forgez, P., Iannelli, A., Bucci, C., Alifano, M., & Alifano, P. (2021). Molecular docking simulation reveals ACE2 polymorphisms that may increase the affinity of ACE2 with the SARS-CoV-2 Spike protein. *Biochimie*, 180, 143-148.

48. Luo, Y., Liu, C., Guan, T., Li, Y., Lai, Y., Li, F., ... & Zeyaweidong, A. (2019). Association of ACE2 genetic polymorphisms with hypertension-related target organ damages in south Xinjiang. *Hypertension Research*, 42(5), 681-689.
49. Bhanushali, A., Rao, P., Raman, V., Kokate, P., Ambekar, A., Mandva, S., ... & Das, B. R. (2018). Status of TMPRSS2–ERG fusion in prostate cancer patients from India: correlation with clinico-pathological details and TMPRSS2 Met160Val polymorphism. *Prostate international*, 6(4), 145-150.
50. Srivastava, A., Bandopadhyay, A., Das, D., Pandey, R. K., Singh, V., Khanam, N., ... & Chaubey, G. (2020). Genetic association of ACE2 rs2285666 polymorphism with COVID-19 spatial distribution in India. *Frontiers in genetics*, 1163.
51. Ravikanth, V., Sasikala, M., Naveen, V., Latha, S. S., Parsa, K. V. L., Vijayasarathy, K., ... & Reddy, D. N. (2021). A variant in TMPRSS2 is associated with decreased disease severity in COVID-19. *Meta Gene*, 29, 100930.
52. Hou, Y., Zhao, J., Martin, W., Kallianpur, A., Chung, M. K., Jehi, L., ... & Cheng, F. (2020). New insights into genetic susceptibility of COVID-19: an ACE2 and TMPRSS2 polymorphism analysis. *BMC medicine*, 18, 1-8.
53. Dobrijevic, Z., Robajac, D., Gligorijevic, N., Šunderic, M., Penezic, A., Miljuš, G., & Nedic, O. (2022). The association of ACE1, ACE2, TMPRSS2, IFITM3 and VDR polymorphisms with COVID-19 severity: A systematic review and meta-analysis. *EXCLI journal*, 21, 818.

54. Gupta, K., Kaur, G., Pathak, T., & Banerjee, I. (2022). Systematic review and meta-analysis of human genetic variants contributing to COVID-19 susceptibility and severity. *Gene*, 146790.
55. Martinez-Gomez, L. E., Herrera-Lopez, B., Martinez-Armenta, C., Ortega-Pena, S., Camacho-Rea, M. D. C., Suarez-Ahedo, C., ... & Lopez-Reyes, A. (2022). ACE and ACE2 gene variants are associated with severe outcomes of COVID-19 in men. *Frontiers in immunology*, 450.
56. Wulandari, L., Hamidah, B., Pakpahan, C., Damayanti, N. S., Kurniati, N. D., Adiatmaja, C. O., ... & Oceandy, D. (2021). Initial study on TMPRSS2 p. Val160Met genetic variant in COVID-19 patients. *Human genomics*, 15(1), 1-9.
57. Saengsiwaritt, W., Jittikoon, J., Chaikledkaew, U., & Udomsinprasert, W. (2022). Genetic polymorphisms of ACE1, ACE2, and TMPRSS2 associated with COVID-19 severity: A systematic review with meta-analysis. *Reviews in medical virology*, 32(4), e2323.
58. Sama, I. E., Ravera, A., Santema, B. T., Van Goor, H., Ter Maaten, J. M., Cleland, J. G., ... & Voors, A. A. (2020). Circulating plasma concentrations of angiotensin-converting enzyme 2 in men and women with heart failure and effects of renin–angiotensin–aldosterone inhibitors. *European heart journal*, 41(19), 1810-1817.
59. Zhang, H., Penninger, J. M., Li, Y., Zhong, N., & Slutsky, A. S. (2020). Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular

mechanisms and potential therapeutic target. *Intensive care medicine*, 46, 586-590.

60. Cheng, Z., Zhou, J., To, K. K. W., Chu, H., Li, C., Wang, D., ... & Yuen, K. Y. (2015). Identification of TMPRSS2 as a susceptibility gene for severe 2009 pandemic A (H1N1) influenza and A (H7N9) influenza. *The Journal of infectious diseases*, 212(8), 1214-1221.

61. Barash, A., Machluf, Y., Ariel, I., & Dekel, Y. (2020). The pursuit of COVID-19 biomarkers: putting the spotlight on ACE2 and TMPRSS2 regulatory sequences. *Frontiers in medicine*, 7, 582793.

62. Hou, Y., Zhao, J., Martin, W., Kallianpur, A., Chung, M. K., Jehi, L., ... & Cheng, F. (2020). New insights into genetic susceptibility of COVID-19: an ACE2 and TMPRSS2 polymorphism analysis. *BMC medicine*, 18, 1-8.

63. Schönfelder, K., Breuckmann, K., Elsner, C., Dittmer, U., Fistera, D., Herbstreit, F., ... & Möhlendick, B. (2021). Transmembrane serine protease 2 polymorphisms and susceptibility to severe acute respiratory syndrome coronavirus type 2 infection: a German case-control study. *Frontiers in genetics*, 12, 667231.

64. FitzGerald, L. M., Agalliu, I., Johnson, K., Miller, M. A., Kwon, E. M., Hurtado-Coll, A., ... & Huntsman, D. G. (2008). Association of TMPRSS2-ERG gene fusion with clinical characteristics and outcomes: results from a population-based study of prostate cancer. *BMC cancer*, 8, 1-10.

65. Yaghoobi, A., Lord, J. S., Rezaiezhadeh, J. S., Yekaninejad, M. S., Amini, M., & Izadi, P. (2023). TMPRSS2 polymorphism (rs12329760) and the severity of the COVID-19 in Iranian population. *PLoS One*, *18*(2), e0281750.
66. Liu, H., & Hu, Y. (2010). Hardy-Weinberg equilibrium in genetic epidemiology. *Zhong nan da xue xue bao. Yi xue ban= Journal of Central South University. Medical sciences*, *35*(1), 90-93.
67. Garavito, P., Mosquera-Heredia, M. I., Fang, L., Payares, F., Ruiz, M., Arias, I., ... & Silvera-Redondo, C. (2020). Polimorfismos de los genes del sistema leptina-melanocortina asociados con la obesidad en la población adulta de Barranquilla. *Biomédica*, *40*(2), 257.
68. Barroso, I., Gurnell, M., Crowley, V. E. F., Agostini, M., Schwabe, J. W., Soos, M. A., ... & O'Rahilly, S. (1999). Dominant negative mutations in human PPAR γ associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature*, *402*(6764), 880-883.
69. Stephens, J. W., Dhamrait, S. S., Cooper, J. A., Acharya, J., Miller, G. J., Hurel, S. J., & Humphries, S. E. (2005). The D allele of the ACE I/D common gene variant is associated with Type 2 diabetes mellitus in Caucasian subjects. *Molecular genetics and metabolism*, *84*(1), 83-89.
70. Yilmaz, H., Agachan, B., Ergen, A., Karaalib, Z. E., & Isbir, T. (2004). Methylene tetrahydrofolate reductase C677T mutation and left ventricular

hypertrophy in Turkish patients with type II diabetes mellitus. *BMB Reports*, 37(2), 234-238.

71. Vimalleswaran, K. S., Radha, V., & Mohan, V. (2006). Thr54 allele carriers of the Ala54Thr variant of FABP2 gene have associations with metabolic syndrome and hypertriglyceridemia in urban South Indians. *Metabolism*, 55(9), 1222-1226.

72. Sanghera, D. K., Ortega, L., Han, S., Singh, J., Ralhan, S. K., Wander, G. S., ... & Kamboh, M. I. (2008). Impact of nine common type 2 diabetes risk polymorphisms in Asian Indian Sikhs: PPARG2 (Pro12Ala), IGF2BP2, TCF7L2 and FTO variants confer a significant risk. *BMC medical genetics*, 9(1), 1-9.

73. Baudin, B. (2005). Polymorphism in angiotensin II receptor genes and hypertension. *Experimental physiology*, 90(3), 277-282.

13. ANEXOS

13.1. Anexo 1. Resumen proyecto “Respuesta Inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile” y consentimiento informado.



Consentimiento Informado

TÍTULO: Análisis de la respuesta inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile

Investigadores

Nombre Investigador responsable (IR) Coordinadora en región Biobío
Dra. LILIANA LAMPERTI, Universidad de Concepción. (llampert@udec.cl)
Investigador Principal (IP) en el centro Instituto Salud Pública ISP
Dr. ABEL VASQUEZ, Sección Virología del ISP de Chile. (avasquez@ispch.cl)

Nombre Co. investigadores locales

Dr. Felipe Zúñiga, Docente del Dpto. de Bioquímica Clínica e Inmunología de Facultad de Farmacia. Universidad de Concepción.
Dra. Estefanía Nova, Docente del Dpto. de Bioquímica Clínica e Inmunología de Facultad de Farmacia. Universidad de Concepción.
Mg. Daniela Catillo, Docente del Dpto. de Bioquímica Clínica e Inmunología de Facultad de Farmacia. Universidad de Concepción.
Dra. Valeska Ormazabal, Docente del Dpto. de Farmacología, Facultad de Ciencias Biológicas. Universidad de Concepción.

Estimado paciente:

Este es un estudio para evaluar la respuesta inmune de la población chilena vacunada contra SARS-CoV-2, así como en individuos recuperados de COVID-19. Por favor lea este documento de consentimiento con tranquilidad, se le dará tiempo para conversar con el profesional a cargo, quien le explicará el contenido de este documento y resolverá sus dudas.

Fundamentos:

El virus SARS-CoV-2 es el causante de la enfermedad COVID-19, patología altamente infectiva que afecta a un amplio espectro de la población, siendo más susceptibles los adultos mayores y aquellas personas que presentan factores de riesgo como obesidad, hipertensión o diabetes principalmente. Esta enfermedad tiene una tasa de letalidad entre el 2-3%, y aproximadamente el 10% de los infectados cursa el cuadro severo de la enfermedad. Los datos de la literatura describen que cerca del 25% de los infectados son asintomáticos, siendo estos una de las principales causas de propagación del virus.

En Chile se aplican diferentes vacunas contra SARS-CoV-2 y el objetivo de este estudio es evaluar la respuesta inmune que las vacunas (Pfizer y Sinovac) o la infección viral en casos graves, generarán en la población chilena.



¿Cuáles son los beneficios de evaluar la respuesta inmune inducida por las diferentes vacunas contra SARS-CoV-2 en la población chilena?

Generar información importante sobre la respuesta inmunológica a la vacunación en la población chilena, permitiendo evaluar la protección a la infección inducida por las vacunas en la población incluida en este estudio.

¿En qué consiste?

Se evaluará la respuesta del sistema inmune en personas vacunadas. Además, se evaluará la capacidad neutralizante de los anticuerpos inducida por vacunación contra la infección viral.

¿Cómo lo vamos a hacer?

Si Ud. acepta participar de manera libre y voluntaria y firma el consentimiento informado, profesionales capacitados del Cesfam para tales efectos realizarán los siguientes procedimientos:

1. Toma de una muestra de hisopado nasofaríngeo o saliva (optativo) para analizar la presencia de SARS-CoV-2 por RT-PCR.
2. Toma de una muestra única de 20 mL de sangre, equivalente a 2 cucharadas soperas (4 tubos de 5 mL). Donde una se utilizará para obtener suero y las otros tres para obtener plasma y linfocitos T.

¿Qué haremos con sus muestras?

Las muestras obtenidas en la comuna de Concepción, en la Región del Biobío, serán trasladadas a la Universidad de Concepción, donde se analizará las muestras de hisopado nasofaríngeo o saliva para la detección de SARS-CoV-2 y se procesará las muestras de sangre para la obtención de sueros y linfocitos T. En los linfocitos se realizarán detección de variantes para receptor del virus en células humanas. Los linfocitos serán almacenadas a -80 °C hasta su traslado en cadena de frío (nitrógeno líquido y hielo seco) a la Sección Biotecnología del Instituto de Salud Pública de Chile (Santiago de Chile).

El suero será utilizado para evaluar anticuerpos y capacidad neutralizante de la infección por SARS-CoV-2, para lo que se utilizará la secuencia de la proteína S aislada en China al principio de la pandemia. Posteriormente, y cuando se disponga de lo necesario para evaluar la capacidad neutralizante contra las otras variantes del virus de SARS-CoV-2 detectadas en Chile, se utilizarán estos sueros para evaluar la capacidad neutralizante de los anticuerpos inducidos por vacunación contra estas variantes. Mientras que, el plasma será utilizado para caracterizar la respuesta inmune celular.



¿Qué riesgos tiene participar?

No presenta riesgos graves para su salud, se obtendrá una muestra de hisopado nasofaríngeo, de la misma forma que se utiliza para la detección de SARS-CoV-2 por PCR en Tiempo Real. Además, se obtendrá una muestra de sangre similar a la que se obtiene al hacer exámenes de sangre de rutina.

En el caso de hisopado nasofaríngeo, podría sentir una leve molestia de cosquilleo o sensación de estornudo por el roce de la tórula de algodón en la mucosa nasal. Puede quedar con sensación de molestia por unos minutos, lo que pasará y no sentirá mayor incomodidad. Deberá indicar si tiene alguna desviación de tabique o irritación nasal para optar por la muestra de saliva.

En el caso de la toma de muestra venosa de sangre, ciertos riesgos en algunos casos, como la sensación de mareo, un leve sangramiento que deberá pasar dentro de los primeros minutos y en muy pocos casos existe riesgo leve de formar un pequeño hematoma (moretón) que desaparecerá en dos días aproximadamente. Puede quedar una sensación de dolor que pasará en los primeros minutos. Riesgos personalizados: se consultará si el individuo(a) está con tratamiento anticoagulante para evitar riesgos de sangramiento y en aquellos casos se observará por más de 5 minutos la formación del tapón plaquetario y cese de sangrado en el brazo del paciente. Sujetos que indiquen desmayos anteriores, serán tratados con mayor cuidado con personal de asistencia en caso de un eventual desmayo.

¿Quiénes no pueden participar?

Personas:

- Fuera del rango etario exigido (menor de 18 o mayor de 80 años)
- Con COVID-19 activo/a
- Que no firmen el consentimiento informado
- Con esquema de inmunización incompleto, falta de alguna dosis de la vacuna Sinovac (dos dosis) o vacuna Pfizer (1 dosis).
- Con última dosis de vacuna recibida hace menos de 21 días.

¿Qué sucede si no acepta participar?

Nada, su participación es completamente libre y voluntaria.



¿Qué haremos con los datos que obtengamos de su participación?

Se analizarán estadísticamente para evaluar la respuesta inmune estimulada por la vacunación en cada participante de este estudio. En ningún caso se utilizarán los nombres de los participantes. Todos los datos serán procesados de forma anónima. La información inmunológica obtenida en este estudio podría ser publicada en revistas científicas nacionales o internacionales, y en un informe de circulación institucional y del Ministerio de Salud.

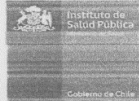
Contacto en caso de consultas:

Investigador principal del ISP Dr. Abel Vásquez Veloso, ubicable en el teléfono 225755513 e-mail: avasquez@ispch.cl de Santiago

Investigador responsable, coordinadora región Biobío: Dra. Liliana Lamperti Fernández en el teléfono +56976343016 e-mail: llampert@udec.cl de Concepción.

Presidente Comité Ético Científico SS Concepción Dr. Nelson Pérez T. Fono 41-2722745 (Para consultas sobre sus derechos)

Presidenta Comité de Ética, Bioética y Bioseguridad de la Dirección de Investigación y Creación Artística de la Vicerrectoría de Investigación y Desarrollo de la Universidad de Concepción, Dra. María Andrea Rodríguez Tastets, teléfono de contacto 41-2204302.



Consentimiento Informado

TÍTULO: Análisis de la respuesta inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile

DECLARACIÓN DE CONSENTIMIENTO:

Si usted ha entendido los objetivos y lo que le estamos solicitando, ha hecho todas las preguntas necesarias y acepta voluntariamente participar, por favor firme este consentimiento en el cual Ud. declara que:

- He leído este formulario.
- He podido realizar todas mis consultas.
- He recibido respuestas a mis dudas.
- He comprendido el contenido de este formulario.
- Certifico que el procedimiento a realizar se me ha explicado en detalle.

Nombre del participante (escrito por participante) firma fecha

Nombre de la persona que solicita el consentimiento firma fecha

VILMA RAZMILIC BONAZIC

Nombre del director de la institución (o delegado) firma fecha

13.2. Anexo 2. Documento de aprobación del Comité Ética Científico Concepción.



ACTA DE EVALUACIÓN COMITÉ ÉTICO CIENTÍFICO

COMITÉ CONSTITUIDO EN CONFORMIDAD A LA R. EX N°2444 / 14.08.2012
RE ACREDITADO POR LA SEREMI SALUD DEL BIOBIO EN CONFORMIDAD A LA R. EX. N° 2963/ 06.10.2020

Concepción, 28 de octubre 2021

Dra. Liliana Lamperti
Investigadora Responsable
Dpto. Bioquímica Clínica e Inmunología
Facultad de Farmacia
Universidad de Concepción
Presente.

Ref: Análisis de la respuesta inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile
Código CEC-SSC: 21-09-47

Estimado investigador:

Junto con saludarle, nos permitimos informar a usted que el Comité Ético Científico del Servicio de Salud Concepción, recepcionó el 26 de octubre 2021 las enmiendas efectuadas al protocolo del estudio de la referencia, en respuesta a las observaciones planteadas en el acta de este comité de fecha 05 de octubre 2021. En consideración a los antecedentes presentados, este comité aprueba en forma expedita las correcciones en comento y se deja establecido que este procedimiento será informado en la próxima reunión del comité, a realizarse el 02 de noviembre 2021.

Documentos recepcionados y revisados:

- Carta respuesta observaciones
- Anexo 1 Consentimiento Análisis respuesta inmune vacunados en Chile versión 4 25-10-2021 CON CONTROL DE CAMBIO
- Anexo 1 Consentimiento Análisis respuesta inmune vacunados en Chile versión 4 25-10-2021 SIN CONTROL DE CAMBIO
- Anexo 2 Data participantes Análisis respuesta inmune vacunados en Chile versión 3 25-10-2021
- Anexo 2-1 Data participantes Análisis respuesta inmune vacunados en Chile versión 3 25-10-2021

CONCLUSION: Con fecha 28 de octubre 2021, el Comité **aprueba en forma expedita** el estudio, y procede a timbrar los siguientes documentos:

- Protocolo de investigación: Análisis de la respuesta inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile Versión 3.0/27-09-2021
- Compromisos de toma de muestra, bioseguridad y uso EPP
- Compromisos de traslado de muestras y aseguramiento de procedimiento de toma de muestra
- Anexo 1 Consentimiento Análisis respuesta inmune vacunados en Chile versión 4 25-10-2021 SIN CONTROL DE CAMBIO
- Anexo 2 Data participantes Análisis respuesta inmune vacunados en Chile versión 3 25-10-2021
- Anexo 2-1 Data participantes Análisis respuesta inmune vacunados en Chile versión 3 25-10-2021

EN CASO DE CUALQUIER DUDA SE LE SOLICITA CONTACTARSE CON EL CEC DEL SSC. Página 1
ESTE COMITÉ CUMPLE CON LAS GUÍAS DE BUENA PRÁCTICA CLÍNICA DEFINIDAS POR LA CONFERENCIA INTERNACIONAL DE ARMONIZACIÓN (GCP-ICH); CON LAS LEYES CHILENAS N° 20.120 Y N° 19.628 Y CON LAS NORMAS INSTITUCIONALES DEL SERVICIO DE SALUD CONCEPCIÓN



ACTA DE EVALUACIÓN COMITÉ ÉTICO CIENTÍFICO

COMITÉ CONSTITUIDO EN CONFORMIDAD A LA R. EX N°2444 / 14.08.2012
RE ACREDITADO POR LA SEREMI SALUD DEL BIOBIO EN CONFORMIDAD A LA R. EX. N° 2963/ 06.10.2020

Además, este Comité solicita que:

1. El investigador utilice el formulario de consentimiento informado con el timbre y fecha de aprobación del comité en los pacientes que serán enrolados en su institución.
2. El Investigador responsable debe cumplir con el adecuado manejo de la información, asegurando la no utilización de datos que pueda identificar a los participantes.
3. Las enmiendas, reportes de eventos adversos y desvíos de protocolo se entreguen al comité en conjunto con una opinión personal del investigador, sobre la relevancia de ellos en el proyecto.
4. Informar a lo menos una vez al año de su estado de desarrollo, como también de cualquier publicación o presentación a congresos que de él se generen
5. La validez de esta aprobación es por un año.
6. Una vez finalizado el proyecto, el comité deberá ser informado de los resultados del estudio.
7. Conservar toda la documentación en su poder por lo menos hasta quince años cerrado el estudio
8. En toda futura correspondencia hacer referencia al número de código asignado
9. Cumplir con la normativa y protocolos de seguridad del establecimiento y/o centro de investigación

Atentamente,


Ing. Ricardo Saldías Cerda
Secretario CEC-SSC

CC: CEC-SSC



Firmado con firma electrónica
Avanzado por
NELSON GOR PEREZ TERAN
Fecha 2021.10.29 10:21:17 -0300

Dr. Nelson Perez Terán
Presidente CEC-SSC

13.3. Anexo 3. Cálculo tamaño muestral para el proyecto “Respuesta inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile”.

Se consideró la inmunogenicidad reportada por los fabricantes de la vacuna BNT162b2 de Pfizer-BioNTech (35) y CoronaVac® de SINOVAR (33) para estimar el tamaño de muestra mínimo requerido y, de manera análoga, se consideró la memoria inmunológica de las células del sistema inmune adaptativo de post-infectados (36). El tamaño de muestra se estimó con un 95% de confianza, un 4% de error y se ajustó con un 20% de pérdida (Ecuación 1). El tamaño muestral final se distribuirá mediante afijación proporcional y posteriormente por cada estrato de sexo y grupo etario.

Se utilizó la Ecuación 1 para calcular el tamaño de muestra mínimo requerido para estimar la proporción de participantes con respuesta inmune protectora respecto a SARS-CoV-2 para cada vacuna y el grupo de post-infectados.

$$n_{pre} = \frac{z_{\alpha/2}^2 * P * (1-P)}{d^2} \quad (\text{Ecuación 1})$$

Donde:

z : es el percentil de la función de distribución de la Gaussiana (Normal (0,1)). Este valor corresponde a un 1,96 dado que se está considerando un 95% de confianza.

α : es la probabilidad que el intervalo de confianza no contenga al parámetro. Este valor de significancia corresponde a un 0,05 dado que se está considerando un 95% de confianza.

P : es el porcentaje reportado de inmunogenicidad en vacunados o post-infectados (ver Tabla 1).

d : es el error máximo admisible, valor que corresponde a un error de tolerancia para la estimación del parámetro. Este valor corresponde a un 0,04 dado que se está considerando un error del 4% entorno a la estimación del parámetro. Los parámetros iniciales para P se encuentran en la Tabla 1.

El tamaño muestral final se ajustó por pérdida mediante la Ecuación 2.

$$n_{final} = \frac{n_{pre}}{1-R} \quad (\text{Ecuación 2})$$

donde:

R : es la proporción de pérdida esperada

Tabla 1: Inmunogenicidad reportada según origen de la muestra.

Origen de la Muestra	Inmunogenicidad
Vacunados con Pfizer-BioNTech	91,9%*
Vacunados con Coronavac	90,7%†
No vacunados (post-infectados)	10%‡

*: Inmunogenicidad reportado en Sahin et al., 2020

†: Inmunogenicidad reportado en Wu et al., 2021

‡: Inmunogenicidad reportado en Rodda et al., 2021

El tamaño de muestra total corresponde a 749 pacientes desglosados según indica la Tabla 2.

Tabla 2: Tamaño muestral total según origen de la muestra.

Origen de la Muestra	Tamaño muestral total
Vacunados con Pfizer-BioNTech	223
Vacunados con Coronavac	254
No vacunados (post-infectados)	272
Total	749

El tamaño muestral para cada región será de un 50% del total y ajustable según origen de la muestra, en especial para el caso de NO vacunado (post -infección)

13.4. Anexo 4. Formulario de datos autoreportados por los voluntarios del proyecto “Respuesta inmune en individuos recuperados de COVID-19 y vacunados contra SARS-CoV-2 en Chile”.



Formulario

TOMA DE MUESTRA Y LABORATORIO
COMPLETAR POR EL PROFESIONAL RESPONSABLE DE TOMA DE MUESTRA

DATOS PERSONALES DEL PARTICIPANTE

Rut: _____ Teléfono: _____

Correo electrónico: _____

Nombre: _____ Apellido paterno: _____

Apellido materno: _____

Fecha Nacimiento: ____/____/____ Edad: ____

Género: M __ F __ Otros __

Embarazo: SI: _____ NO: _____

Tiempo de gestación (semanas): _____

Peso (kg): _____ Talla (cm): _____ (datos autoreportados)

Región: ____ Comuna: _____

Servicio de Salud: _____

Establecimiento: _____

Antecedentes de Morbilidad:

Diabetes: SI: _____ NO: _____

HTA: SI: _____ NO: _____

Obesidad: SI: _____ NO: _____

Inmunosuprimida: SI: _____ NO: _____

Diagnóstico de inmunosupresión: _____

Tratamiento de inmunosupresión: _____

Fue contacto estrecho por COVID-19: SI: _____ NO: _____ FECHA: _____

Estuvo enfermo por COVID-19: SI: _____ NO: _____

- Fecha de infección por COVID-19: _____

- Hospitalizado por COVID-19: SI: _____ NO: _____



- Fecha de ingreso: _____ Fecha de egreso: _____
- Carnet de Vacunación a la vista: SI: _____ NO: _____
- Vacuna: Pfizer/BioNTech: _____ ó CoronaVac®/Sinovac: _____
- Fecha de primera dosis: _____
- Fecha de segunda dosis: _____ (para Sinovac)
- Sintomatología COVID-19: SI: _____ NO: _____

Fecha de toma de muestra de HNF y de sangre: ____/____/____

Hora de toma de muestra: _____

Profesional responsable de toma de muestra: _____

Completar por laboratorio Universitario Concepción

Fecha de recepción: _____ Hora de recepción: _____

Responsable de ingreso de muestra: _____

n° de muestra: _____

Responsable de procesamiento: _____

Fecha de entrega al ISP: _____

Observaciones:

Recepción de muestra en el ISP

Fecha de recepción: _____ Hora de recepción: _____

Responsable de ingreso de muestra: _____

n° de muestra: _____

Responsable de procesamiento: _____

Observaciones:

13.5. Anexo 5. Clasificación y manejo de residuos biológicos.

CLASIFICACIÓN Y MANEJO DE RESIDUOS BIOLÓGICOS

Tipo de Residuo	Definición	Almacenamiento y Retiro	Contenedor o bolsa
Residuos biológicos (tipo 1)	Corresponden SOLO a tejidos humanos, restos de tejidos humanos, materiales contaminados con fluidos humanos (sangre, orina) y restos de animales contaminados con bacterias o virus patógenos.	Deben ser almacenados a temperaturas inferiores a 4°C hasta ser retirados por la unidad RESPEL o empresa externa*.	Bolsa de autoclave amarilla 
Residuos cortopunzantes biológicos (tipo 2)	Corresponden a materiales cortopunzantes (agujas, lancetas, hojas de bisturí, escalpelos) que se han utilizado con fluidos y tejidos humanos o en disecciones de animales contaminados con bacterias o virus patógenos.	Deben ser almacenados a temperaturas inferiores a 4°C hasta ser retirados por la unidad RESPEL o empresa externa*.	Caja cortopunzante amarilla 
Residuos bio-inertes (tipo 3)	Corresponden a animales o restos de animales de experimentación que NO están contaminados con bacterias o virus patógenos u otras sustancias peligrosas.	Deben ser almacenados en bolsas para residuos inertes color negro, resistente al volumen y peso del residuo, a temperaturas inferiores a 4°C hasta ser retiradas por la unidad RESPEL para su disposición final. En caso de que el residuo corresponda a una muestra preservada, se debe separar de la fracción líquida (ej. formalina, alcohol), la que debe ser almacenada como residuo químico.	Bolsa para residuos inertes color negro 
Residuos cortopunzantes inertes (tipo 4)	Corresponden a materiales cortopunzantes (agujas, lancetas, hojas de bisturí, escalpelos) que NO han estado en contacto con tejidos humanos, restos de tejidos humanos, fluidos humanos (sangre, orina), ni restos de animales contaminados con bacterias o virus patógenos.	Deben ser almacenados en contenedores plásticos de 4 u 8 kg, etiquetados como "Cortopunzante inerte", hasta ser retirados por la unidad RESPEL para su disposición final	Contenedor plástico 4 u 8 kg 
Residuos microbiológicos (tipo 5)	Corresponden a materiales de cultivo que han estado en contacto con bacterias no patógenas y que han sido inactivados por esterilización con autoclave o por desinfección en un baño con solución de hipoclorito de sodio.	Deben ser entregados a la unidad RESPEL, SOLO después de ser inertizados mediante autoclave o desinfectados en un baño con solución de hipoclorito de sodio al 1% durante 1 hora. En caso contrario, no se procederá con el retiro. La unidad RESPEL entregará bolsas de autoclave solo cuando éste sea el método utilizado de inactivación.	Bolsa de autoclave transparente o blanca (solo para autoclave) 

(*) El retiro desde la Facultad de Medicina, Facultad de Odontología, DISE y Centro RAI, será una vez a la semana a través de empresa externa Stericycle y la unidad RESPEL. El resto de las facultades que generen este tipo de residuos deberá gestionar retiro mensual a través de la unidad RESPEL.