



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO

---

---

FACULTAD DE CIENCIAS

PREDICCIÓN DE EPÍTOPOS  
DE CÉLULAS B Y T DE SARS-CoV-2  
A TRAVÉS DE ANÁLISIS BIOINFORMÁTICO

**T E S I S**

QUE PARA OBTENER EL TÍTULO DE:

BIÓLOGA

P R E S E N T A:

MARÍA GUADALUPE GARCÍA VALERIANO



**DIRECTORES DE TESIS:**  
DRA. TANIA RIVERA HERNÁNDEZ  
DR. CONSTANTINO III ROBERTO LÓPEZ  
MACÍAS

2022

Ciudad Universitaria, CD. MX.,



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

FACULTAD DE CIENCIAS  
Secretaría General  
División de Estudios Profesionales

Votos Aprobatorios

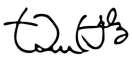
LIC. IVONNE RAMÍREZ WENCE  
Directora General  
Dirección General de Administración Escolar  
Presente

Por este medio hacemos de su conocimiento que hemos revisado el trabajo escrito titulado:

**Predicción de epítomos de células B y T de SARS-CoV-2 a través de análisis bioinformático**

realizado por **María Guadalupe García Valeriano** con número de cuenta **313103285** quien ha decidido titularse mediante la opción de **tesis** en la licenciatura en **Biología**. Dicho trabajo cuenta con nuestro voto aprobatorio.

Propietario Dr. Luis Alfonso Vaca Domínguez  
Propietario Dr. Luis David Alcaraz Peraza  
Propietaria Dra. Tania Rivera Hernández  
Tutora  
Suplente Dra. Margarita Jacaranda Rosendo Pineda  
Suplente Dr. Constantino III Roberto López Macías



Atentamente  
"POR MI RAZA HABLARÁ EL ESPÍRITU"  
CIUDAD UNIVERSITARIA, CD. MX., A 27 DE OCTUBRE DE 2021

JEFE DE LA DIVISIÓN DE ESTUDIOS PROFESIONALES

ACT. MAURICIO AGUILAR GONZÁLEZ

Señor sinodal: antes de firmar este documento, solicite al estudiante que le muestre la versión digital de su trabajo y verifique que la misma incluya todas las observaciones y correcciones que usted hizo sobre el mismo.

*A mis abuelos María Teresa Francisco y José Ángel Valeriano,  
cada letra aquí plasmada es un avance y ese avance es por y para ustedes.*

*Abuelita, gracias por todas tus enseñanzas, las llevo  
conmigo para toda la vida y sé que desde donde te encuentras  
iluminas mis pasos para nunca dejarme sola.*

*Abuelito, gracias por no permitirme darme por vencida, por  
tus sabias palabras y por tu apoyo incondicional.*

*Gracias por ser mi inspiración y mis mayores ejemplos de vida.  
Los amo.*

## **AGRADECIMIENTOS**

Gracias a la Universidad Nacional Autónoma de México y a la Facultad de Ciencias, por todo lo aprendido dentro de sus aulas y por todas las oportunidades que me ha brindado.

Gracias a mis sinodales el Dr. Luis Vaca, la Dra. Margarita Rosendo y el Dr. Luis Alcaraz, los tres fueron maestros extraordinarios. Espero algún día marcar la vida de alguien como cada uno de ustedes marcó la mía.

Gracias interminables a mi tutora, mentora y siempre modelo a seguir, la Dra. Tania Rivera Hernández por su ayuda, su comprensión y su infinita paciencia. Qué bonito fue aprender de usted. Estaré eternamente en deuda por el incuantificable impulso que le ha dado a mi carrera. Es la mejor tutora y una excelente persona.

Gracias infinitas a mi tutor, el Dr. Constantino López Macías por confiar en mí y recibirme en su laboratorio, porque encontré en usted guía y apoyo. Que orgullo culminar bajo su mentoría, ¡gracias!

Gracias a mis maestras y amigas, M. en C. Diana Laura Pacheco Olvera y M. en C. Stephanie Saint Remy Hernández, ¡lo hemos logrado! Gracias por permitirme aprender de su lado, por guiarme paso a paso en la realización de este proyecto, por su paciencia, por resolver mis dudas, por sus múltiples consejos y porque pudiera cumplir una de mis metas más anheladas. ¡Son las mejores científicas!

Gracias especiales al Dr. Luis Alberto Ontiveros Padilla, por asesorarme y compartir sus conocimientos desde la clase de inmunología en la Facultad de Química, por confiar en mí y por la huella que has dejado en mis aspiraciones.

Gracias a todos mis amigos de la UIMIQ por recibirme en su equipo y por enseñarme tanto sobre el amor a la ciencia. A Gonzalo, Vanessa, Jessica, Alejandro, Isauro, Dariana, Andrea y Lety.

Gracias a mis padres Marcela Valeriano y Ramón García. Mamá, gracias por apoyarme siempre sin dudar, por ser mi motor, alentarme y jamás soltarme. Papá, gracias por guiar mi camino, por tu apoyo y por siempre cumplir todos mis sueños. A ambos, infinitas y eternas gracias por dárme todo, por nunca dejar de creer en mí y por su amor, espero que algún día pueda retribuirles todo lo que ustedes me han dado.

Gracias de manera especial a mi hermano Alan García. Gracias por creer en mí, por aterrizarme, cuidarme, acompañarme y por no dejar de recordarme las cosas verdaderamente importantes de la vida. Te admiro y te amo cada día más, juntos siempre hermanable.

Gracias a Josefina Valeriano por ser otra madre para mí, por estar incondicionalmente, por tu complicidad y por cada experiencia que hemos vivido juntas. Eres mi más claro ejemplo de fortaleza, perseverancia y dedicación. Gracias a mi otro hermano Miguel Valeriano por las aventuras compartidas desde niños y por todos esos momentos de tanta felicidad. Eres luz en mi vida y confío en que serás el mejor veterinario del mundo.

Gracias a Jesús Jiménez. No me creerías lo agradecida que estoy con la vida por encontrarte. Gracias por tu tiempo, por creer en mí, por lo feliz que soy de compartir mi vida contigo, por darme calma en los momentos más difíciles y por tu amor sincero e incondicional, te amo de vuelta.

Gracias a mis amigos y biólogos favoritos de la carrera, especialmente a Chema Herrera por ser mi confidente, por la sinceridad con la que me aconsejó en tantos aspectos de la vida y estar incondicionalmente. A Pavel Rodríguez por dejarme

soñar a su lado, por ser el mejor equipo “Marpave”, por nuestros triunfos juntos, y por ser mi mano derecha ante cualquier circunstancia. Gracias a Alejandra Reyes y Daniela Vejar. Amigas mías, tengo tanto que agradecerles que estas líneas no bastarían, gracias por las porras y la escucha. Admiro a las grandes mujeres que se han convertido y deseo innumerables aventuras juntas. Gracias a Mónica Epigmenio porque no puedo ni contar las horas que me regaló para platicar, reír y descargar frustraciones. Gracias por ser una de mis mayores confidentes, por tu complicidad, y por estar en cada paso que di. Gracias a Erik Rodríguez por mostrarme la pasión de amar lo que se hace. Gracias a Itzel Parada por su apoyo, su honestidad y ser una persona increíble. Gracias a Álvaro Santos por permitirme aprender de su lado y el ánimo constante en el camino. Gracias a Jonathan Colín, Aldo Vargas, Baruch Cisneros, Andrea González, Paulina Valverde, Ana Cruz y Jessica Torres, siempre cercanos y disponibles para darme su mano cuando los necesito. Sin ustedes nada hubiera sido lo mismo. Gracias a todos por el acompañamiento y a pesar de que el tiempo y la distancia nos separen, siempre atesoraré los momentos que disfruté a su lado. Los llevo a cada uno en mi corazón.

*“Es casi imposible corresponder a todas las personas que nos ayudan a lo largo de la vida. Quizá sea más sencillo rendirse ante el milagroso alcance de la generosidad humana y seguir diciendo gracias, eterna y sinceramente, mientras nos alcance la voz.”*

-Elizabeth Gilbert

## CONTENIDO

<b>AGRADECIMIENTOS</b> .....	3
<b>ÍNDICE DE FIGURAS Y TABLAS</b> .....	8
<b>RESUMEN</b> .....	11
<b>INTRODUCCIÓN</b> .....	12
<b>ANTECEDENTES</b> .....	14
1.    Coronavirus humanos .....	14
1.1.    Origen y evolución de SARS-CoV-2.....	15
1.2.    Estructura y organización genómica de SARS-CoV-2 .....	16
Glicoproteína Espícula (S) .....	17
Proteína de Membrana (M) .....	20
Proteína de Envoltura (E).....	20
Proteína Nucleocápside (N) .....	20
Proteínas no estructurales y accesorias .....	21
ORF1ab.....	21
ORF3a.....	21
ORF6.....	22
ORF7a y 7b.....	22
ORF8.....	22
ORF10.....	23
1.3.    Ciclo biológico de SARS-CoV-2 .....	23
2.    Respuesta inmunológica frente a SARS-CoV-2 .....	24
2.1.    Determinante antigénico o epítipo.....	26
2.2.    Reactividad cruzada.....	27
2.3.    Complejo principal de histocompatibilidad (MHC) .....	28
3.    Predicción bioinformática de epítomos de SARS-CoV-2 .....	31
3.1.    Métodos de predicción de epítomos .....	32
3.2.    Predicción de epítomos de células B .....	34
3.3.    Predicción de epítomos de células T.....	34
3.3.1.    Predicción de epítomos de unión de MHC-I y MHC-II.....	35
3.3.2.    TepiTool.....	38
3.3.3.    Inmunogenicidad .....	38
4.    Alineación de secuencias .....	39
4.1.    Clustal Omega.....	40



4.2.	EMBOSS Needle .....	40
	<b>PLANTEAMIENTO DEL PROBLEMA</b> .....	42
	<b>JUSTIFICACIÓN</b> .....	43
	<b>PREGUNTA DE INVESTIGACIÓN</b> .....	44
	<b>HIPÓTESIS</b> .....	44
	<b>OBJETIVOS</b> .....	44
	General .....	44
	Particulares.....	44
	<b>MÉTODO</b> .....	46
	Estrategia general.....	46
	Recuperación de datos.....	46
	Predicción de epítomos lineales de células B .....	47
	Predicción de epítomos de células T.....	47
	Alineamiento múltiple.....	49
	Modelado estructural .....	50
	<b>RESULTADOS</b> .....	52
	Predicción de epítomos lineales de células B .....	52
	Predicción de epítomos de células T.....	59
	<b>DISCUSIÓN</b> .....	80
	<b>CONCLUSIÓN</b> .....	91
	<b>PERSPECTIVAS</b> .....	92
	<b>BIBLIOGRAFÍA</b> .....	93
	<b>ANEXOS</b> .....	108
	Anexo 1.....	108
	Anexo 2.....	127
	Anexo 3.....	136

## ÍNDICE DE FIGURAS Y TABLAS

<b>Figura 1.</b> Representación esquemática de SARS-CoV-2.....	16
<b>Figura 2.</b> Organización genética de SARS-CoV-2 .....	17
<b>Figura 3.</b> Estructura del monómero de la proteína S del virus SARS-CoV-2 .....	19
<b>Figura 4.</b> Representación esquemática de los dos tipos de epítomos presentes en los antígenos. ....	27
<b>Figura 5.</b> Mapa de los loci de MHC del cromosoma 6 en el ser humano.....	28
<b>Figura 6.</b> Diagrama de las moléculas de MHC clase I y clase II. ....	30
<b>Figura 7.</b> Diagrama de flujo de las herramientas de predicción disponibles en el recurso IEDB. ....	39
<b>Tabla 1.</b> Número de acceso de las proteínas de SARS-CoV-2 y de los coronavirus humanos.....	47
<b>Tabla 2.</b> Frecuencia en la población mundial de los alelos clase I y II .....	49
<b>Figura 8.</b> Flujo de trabajo del estudio. ....	51
<b>Tabla 3.</b> Epítomos reportados en la literatura de células B lineales. ....	53
<b>Tabla 4.</b> Epítomos de células B conservados, con identidad superior al 50% entre el SARS-CoV-2 y seis coronavirus humanos.....	56
<b>Figura 9.</b> Epítomos lineales de células B .....	58
<b>Tabla 5.</b> Epítomos reportados en la literatura de células T restringidos a MHC-I, MHC-II y promiscuos.....	60
<b>Tabla 6.</b> Epítomos conservados de células T con restricción a MHC-I, MHC-II y promiscuos con identidad superior al 50% entre el SARS-CoV-2 y seis coronavirus humanos.....	73
<b>Figura 10.</b> Epítomos de células T con restricción a MHC-I y promiscuos .....	77
<b>Figura 11.</b> Epítomos de células T con restricción a MHC-II y promiscuos.....	79
<b>Tabla 7.</b> Nuevos epítomos lineales de células B predichos en este estudio. Anexo 1 .....	108
<b>Tabla 8.</b> Nuevos epítomos de células T restringidos a MHC-I, MHC-II y promiscuos predichos en este estudio. Anexo 1. ....	112
<b>Tabla 9.</b> Epítomos reportados experimentalmente y capaces de ser reconocidos para montar una respuesta inmunitaria. Anexo 1. ....	126
<b>Tabla 10.</b> Epítomos conservados de células B con el porcentaje de identidad entre la secuencia de SARS-CoV-2 y seis coronavirus humanos. Anexo 2. ....	127
<b>Tabla 11.</b> Epítomos conservados de células T con restricción MHC-I, MHC-II y promiscuos con el porcentaje de identidad entre la secuencia de SARS-CoV-2 y seis coronavirus humanos. Anexo 2.....	130
<b>Figura 12.</b> <i>Pre-print</i> del manuscrito .....	136

## ABREVIATURAS

H-CoV	Coronavirus humano
SARS-CoV	Síndrome respiratorio agudo severo
SARS-CoV-2	Coronavirus tipo 2 del Síndrome respiratorio agudo severo
MERS-CoV	Síndrome respiratorio de Medio Oriente
HCoV-NL63	Coronavirus humano NL63
HCoV-229E	Coronavirus humano 229E
HCoV-OC43	Coronavirus humano OC43
HKU1	Coronavirus humano HKU1
RaTG13	Coronavirus de murciélago RaTG13
OMS	Organización Mundial de la Salud
COVID-19	Enfermedad por coronavirus 2019
S	Glicoproteína de espícula
M	Proteína de membrana
N	Proteína de la nucleocápside
E	Proteína de envoltura
RBD	Dominio de unión al receptor
NTD	Dominio N-terminal
ORF	Marco de lectura abierto
UTR	Región no traducida
ACE2	Enzima de angiotensina 2
FP	Péptido de fusión
CH	Hélice central
CD	Dominio de conexión
HR1	Dominio de repetición de heptada 1
HR2	Dominio de repetición de heptada 2
TM	Dominio transmembranal
CT	Cola citoplásmica
nsps	Proteínas no estructurales
ARN	Ácido ribonucleico
CPA	Célula presentadora de antígeno
BCR	Receptor de célula B
TCR	Receptor de célula T
ERGIC	Retículo endoplásmico-Golgi
CDR	Regiones determinantes de complementariedad
HV	Regiones hipervariables
MHC	Complejo principal de histocompatibilidad
HLA	Complejo de antígeno leucocitario humano
IEDB	Recurso de análisis y base de datos de epítopos inmunes
NCBI	Centro Nacional de Información Biotecnológica
UniProt	Recurso Proteico Universal
PDB	Banco de Datos de Proteínas
MSA	Alineamiento múltiple de secuencias
ANN	Red neuronal artificial

SMM	Método de matriz de puntuación
SMMPMBEC	SMM con una matriz de covarianza de energía de unión de péptido
IFN $\gamma$	Interferón gamma
UCI	Unidad de cuidados intensivos
VIH	Virus de la inmunodeficiencia humana
PEDV	Diarrea epidémica porcina
pp	Poliproteínas
nm	Nanómetro
aa	Aminoácidos
kb	Kilobases
kDa	Kilodaltones

## RESUMEN

La emergencia sanitaria por la pandemia de COVID-19 se ha posicionado como un problema de salud pública prioritario, ya que desde diciembre de 2019 y hasta el mes de abril de 2022, ha provocado la muerte de más de seis millones de personas a nivel mundial. A pesar de los múltiples esfuerzos por combatir la emergencia sanitaria, aún falta mucho por saber y es necesario generar información relevante que permita controlar la pandemia eficazmente.

La evidencia experimental sugiere que las respuestas de reactividad cruzada entre coronavirus están presentes en la población, sin embargo, se desconoce si los posibles antígenos compartidos entre diferentes coronavirus podrían proporcionar protección cruzada. Teniendo en cuenta que los coronavirus son patógenos emergentes y representan una amenaza para la salud mundial, se debe impulsar el desarrollo de una vacuna contra el SARS-CoV-2 que idealmente tenga el potencial de proporcionar protección “universal” contra otros coronavirus. En esta tesis se presenta una revisión exhaustiva de los epítomos de células B y T reportados compartidos entre el SARS-CoV-2 y otros coronavirus humanos, además, mediante el uso de herramientas bioinformáticas basadas en la web se predijeron nuevos epítomos de células B y T de los cuales, hasta la fecha, no se han encontrado reportes en la literatura.

Los resultados aquí mostrados y la evidencia experimental que surge constantemente nos permiten respaldar la hipótesis de que las respuestas de reactividad cruzada, particularmente las que provienen de las células T, podrían jugar un papel clave en el control de la infección por SARS-CoV-2. Aunado a lo anterior, confiamos que esta evidencia podría apoyar el estudio de respuestas de reactividad cruzada con el fin de dilucidar su papel en la inmunidad al SARS-CoV-2. Finalmente, se espera que estos hallazgos ayuden al análisis dirigido de las respuestas inmunitarias específicas de antígeno y orienten el diseño de vacunas futuras, con el objetivo de desarrollar una vacuna diseñada racionalmente que pueda proporcionar amplia cobertura contra las enfermedades respiratorias causadas por coronavirus.

## INTRODUCCIÓN

En diciembre de 2019, el gobierno chino emitió una alerta sobre los primeros casos de enfermedad respiratoria causada por un nuevo coronavirus. Durante los 3 años siguientes, el patógeno se identificó, se secuenció y se llamó Coronavirus 2 del síndrome respiratorio agudo severo (SARS-CoV-2)<sup>1</sup>, el cual es el responsable de generar la enfermedad denominada COVID-19. El SARS-CoV-2 se extendió por todo el mundo y hasta abril de 2022 ha provocado la muerte de más de seis millones de personas<sup>2</sup>.

En los últimos 20 años han surgido tres brotes importantes por nuevos coronavirus: el síndrome respiratorio agudo severo (SARS-CoV) en 2002, el síndrome respiratorio de Medio Oriente (MERS-CoV) en 2012 y ahora el SARS-CoV-2, todos ellos evolucionando de su hospedador animal original y adaptándose para causar enfermedades respiratorias graves en los seres humanos<sup>1,3</sup>. La interacción humana con la vida silvestre fue clave en la evolución de estos coronavirus y es altamente probable que en el futuro continúen surgiendo más brotes, lo cual representa una amenaza constante para la salud humana<sup>4</sup>.

Esta pandemia se ha tornado en un desafío para la salud pública global, haciéndose evidente que la única herramienta capaz de controlar la propagación del virus es el desarrollo de una vacuna segura y eficaz. De acuerdo con el *vaccine tracker* del periódico *The New York Times* hay más de 270 vacunas candidatas en desarrollo, de las cuales 109 están en diferentes fases de desarrollo preclínico y clínico, y 24 de éstas ya fueron autorizadas. Dentro de las principales vacunas ya aprobadas se encuentran las desarrolladas por Pfizer-BioNTech, Moderna, Oxford/AstraZeneca y Janssen Pharma, las cuales ya se aplican en varios países incluyendo México<sup>5</sup>.

Para desarrollar una vacuna eficaz es necesario comprender la naturaleza antigénica del objetivo principal, en este proyecto el objetivo es SARS-CoV-2, para el cual la mayoría de las vacunas candidatas ya aprobadas se dirigen a la glicoproteína de espícula (S) cuya función es mediar la entrada viral a la célula huésped por medio de la unión del receptor de la enzima de angiotensina 2 (ACE2)<sup>6</sup>. No obstante, se dispone de información limitada sobre la reactividad cruzada,

proceso por el cual algunos epítomos de un antígeno  $X$  son compartidos por otro antígeno  $Y$ , de modo que una proporción de anticuerpos dirigidos contra  $X$  también reaccionarán contra  $Y$ <sup>7</sup>. Por esta razón es importante conocer la similitud antigénica entre la glicoproteína S del SARS-CoV-2 y las proteínas de otros coronavirus, ya sea que estos provoquen enfermedades estacionales o brotes.

Las respuestas inmunitarias contra las infecciones con brotes o coronavirus estacionales difieren significativamente. La evidencia sugiere que la inmunidad a los coronavirus estacionales es de corta duración y los anticuerpos disminuyen rápidamente después de la infección<sup>8</sup>. Adicionalmente la infección controlada de voluntarios humanos sugiere que la infección no genera inmunidad protectora contra la reinfección<sup>9</sup>. Por otro lado, se ha informado que los pacientes recuperados de SARS-CoV portan anticuerpos circulantes por más de 12 años después de ocurrida la infección, aunque no se investigó la actividad neutralizante de dichos anticuerpos<sup>10</sup>. Estos anticuerpos neutralizantes se han considerado el correlato de la protección después de la vacunación o la infección, por lo que han sido el blanco principal de estudio, mientras que se ha prestado poca atención al estudio de las respuestas de células T y su papel en la protección. Debido a esto, es necesario realizar estudios de investigación que ayuden a identificar el papel potencial de epítomos de células B y T que se encuentren en las proteínas de SARS-CoV-2 para la generación de respuestas inmunitarias de reactividad cruzada contra otros coronavirus humanos. Lo anterior respaldando la importancia de sumar esfuerzos para generar medidas integrales para la prevención de futuros brotes de origen zoonótico.

## ANTECEDENTES

### 1. Coronavirus humanos

Los coronavirus humanos (H-CoV) son miembros del orden *Nidovirales*, familia *Coronaviridae* y de la subfamilia *Orthocoronavirinae*. La subfamilia consta de cuatro géneros: *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* y *Deltacoronavirus*. Los alfacoronavirus y betacoronavirus infectan solo a mamíferos, mientras que los gammacoronavirus y los deltacoronavirus infectan aves y algunos de ellos a mamíferos<sup>11</sup>.

En los últimos 20 años, se presentaron tres brotes importantes por nuevos coronavirus provocando afección respiratoria grave en humanos. El primero de ellos designado como síndrome respiratorio agudo severo (SARS) surgió durante los años 2002 y 2003 en Guangdong, China, prolongándose durante 8 meses y propagándose a diferentes partes del mundo; probablemente transmitido por un mamífero pequeño, una civeta (*Paguma larvata*), siendo hospederos intermediarios de un virus procedente de murciélagos de herradura (*Rhinolophus genus*)<sup>11</sup>. En julio de 2003 ya se había detenido su transmisión en todo el mundo, teniendo un recuento de 8,000 casos y más de 750 muertes<sup>12</sup>.

Diez años después del SARS, otro coronavirus altamente patógeno, denominado como coronavirus del síndrome respiratorio del Medio Oriente (MERS-CoV) surgió en los países del Medio Oriente<sup>11</sup>, transmitiéndose directamente a los humanos por camellos dromedarios (*Camelus dromedarius*) y confirmando hasta el 2019 un total de 2,279 casos, de los cuales 806 fueron mortales<sup>13</sup>.

El actual brote provocado por el síndrome respiratorio agudo grave (SARS-CoV-2) que apareció en Wuhan, China, logró extenderse hasta afectar 210 países alrededor del mundo provocando, hasta la fecha, más de seis millones de muertes acumuladas<sup>2,14</sup>.

En conjunto, existen otras cuatro especies de coronavirus estacionales humanos: HCoV-NL63, HCoV-229E, HCoV-OC43 y HKU1, dos de estos pertenecen al género *Alphacoronavirus* y los otros dos al género *Betacoronavirus*,



respectivamente. Todos inducen enfermedades respiratorias en su mayoría leves, aunque en términos genéticos y biológicos cada uno es diferente<sup>11</sup>.

A tres años de la pandemia de COVID-19 aún no queda claro el origen del SARS-CoV-2 y conforme pasa el tiempo surgen nuevas dudas sobre la evolución de su genoma, la aparición de variantes y las consecuencias en el desarrollo de la pandemia. Enseguida, se explica la hipótesis de lo que se sabe hasta ahora.

### **1.1. Origen y evolución de SARS-CoV-2**

Se ha pensado que el SARS-CoV-2 se trata de un virus artificial, sin embargo, casi el 90% de los coronavirus son considerados zoonóticos, es decir, son agentes infecciosos que se transmiten de forma natural de otros animales a los seres humanos o viceversa<sup>15</sup>; sosteniendo la hipótesis de que la mayoría de los coronavirus en humanos se derivan del reservorio del murciélago<sup>16</sup>.

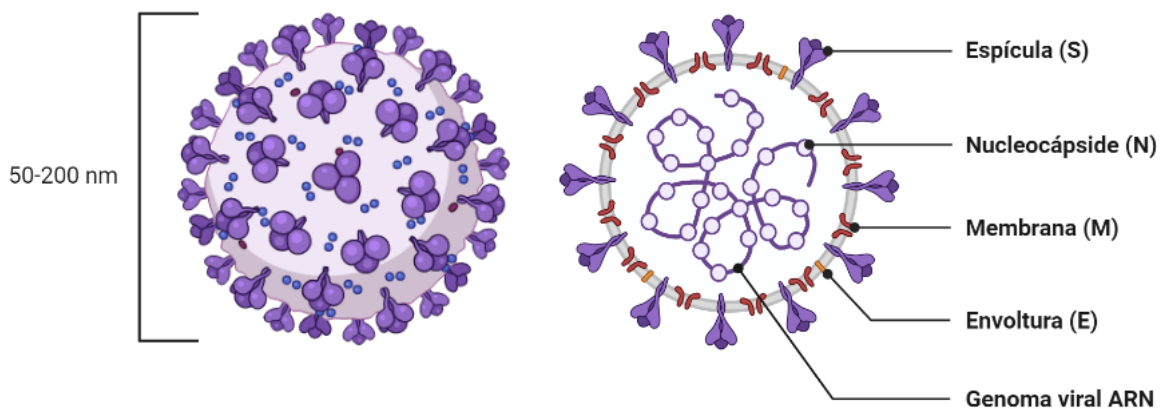
Particularmente, el SARS-CoV-2 está relacionado genéticamente con un betacoronavirus de murciélago del subgénero *Sarbecovirus* y se ha sugerido que infecta al ser humano por medio de un hospedero intermediario: el pangolín malayo (*Manis javanica*), no obstante, esto último aún se encuentra en discusión<sup>17</sup>.

En cuanto a la secuencia del genoma completo del SARS-CoV-2 se ha visto que tiene una similitud del 96.2% con la de un coronavirus denominado RaTG13, encontrado en la especie de murciélagos *Rhinolophus affinis* y presenta aproximadamente un 79% y 50% de similitud con las secuencias del SARS-CoV y el MERS-CoV, respectivamente. En consecuencia, estos datos genéticos orientan a que el SARS-CoV-2 también pudo originarse en murciélagos<sup>17,18</sup>.

Considerando que las poblaciones de virus constantemente evolucionan y presentan varias mutaciones que les podrían conferir alguna ventaja, como la evasión de la respuesta inmune, es importante conocer mejor sus características. En los siguientes párrafos se detallan las características distintivas del SARS-CoV-2 haciendo énfasis en las proteínas que lo conforman y sus funciones.

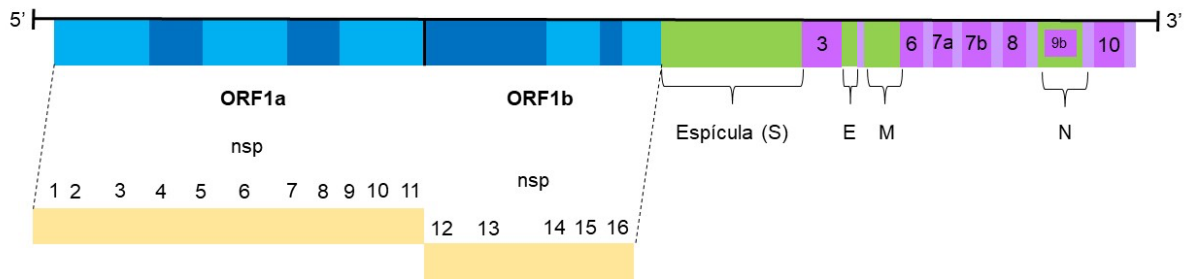
## 1.2. Estructura y organización genómica de SARS-CoV-2

SARS-CoV-2 es un virión de aspecto esférico con un diámetro que varía de 50-200 nanómetros (nm)<sup>19</sup>, contiene un genoma de ARN monocatenario, de polaridad positiva y un tamaño de 26 a 32 kilobases (kb). Su característica más destacada son las proyecciones de picos que emanan de la superficie del virión, dándole la apariencia de una corona, por lo que se le denominó coronavirus (Figura 1)<sup>20</sup>.



**Figura 1. Representación esquemática de SARS-CoV-2.** El virión tiene un diámetro de 50-200 nm cubierto por una membrana lipídica que protege a un genoma de ARN entre 26 y 32 kb. Su genoma es de una sola cadena de ARN, lineal y con sentido positivo que codifica para cuatro proteínas estructurales (S, N, M y E).

El genoma de SARS-CoV-2, que oscila entre 26,000 y 32,000 pares de bases, codifica cuatro proteínas estructurales principales, espícula (S), membrana (M), envoltura (E) y nucleocápside (N)<sup>21</sup>, todas las cuales están codificadas dentro del extremo 3' del genoma viral; y están precedidas por los ORF's (marcos de lectura abierta) que son secuencias de ARN comprendidas entre un codón de inicio de la traducción y un codón de termino, están designados como 1a (pp1a) y 1ab (pp1ab), y codifican para las 16 proteínas no estructurales (nsps). El gen de la replicasa que codifica las nsps ocupa dos tercios del genoma, aproximadamente 20 kb, la parte restante, aproximadamente 10 kb del genoma viral, constituye a las proteínas estructurales y accesorias<sup>18,20,21</sup> (Figura 2).



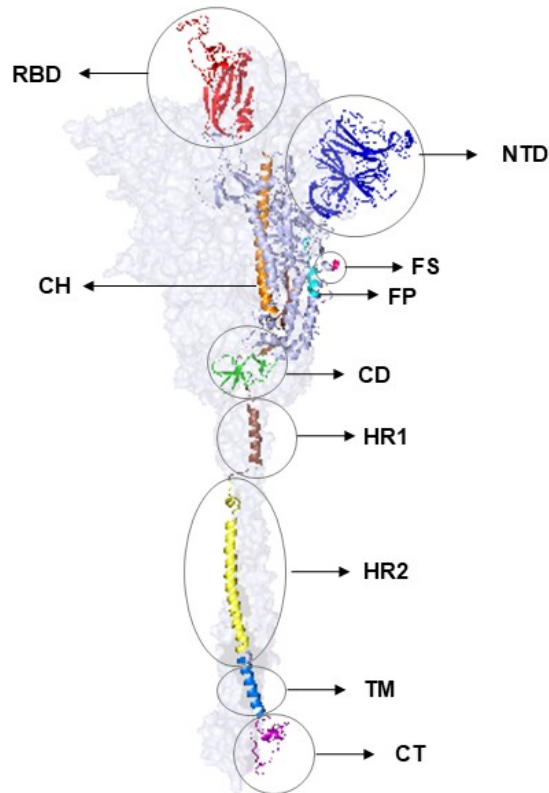
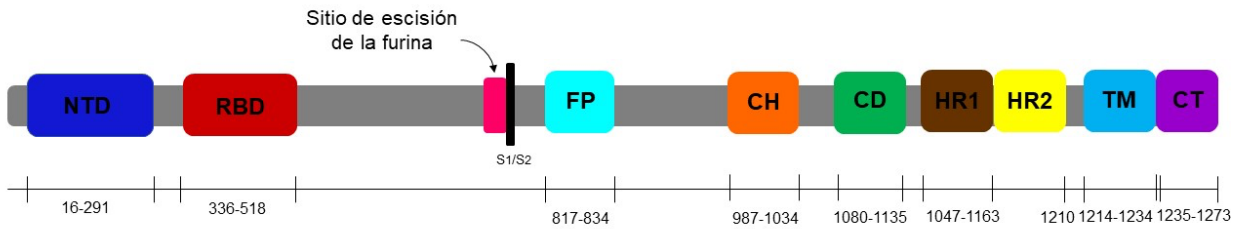
**Figura 2. Organización genética de SARS-CoV-2.** Hacia el extremo 5', se codifica el gen de la replicasa viral por medio de ORF1a y ORF1b. Hacia el extremo 3', se codifican los genes de las 4 proteínas estructurales principales S, E, M y N (en verde) y las de las proteínas accesorias (en morado).

Las cuatro proteínas estructurales principales se describen detalladamente a continuación:

### **Glicoproteína Espícula (S)**

La glicoproteína S es una proteína transmembranal que forma trímeros conformando así a las espículas, las cuales sobresalen hasta 20 nm de la superficie del virión. Su tamaño varía de 1,160 aminoácidos (aa) a 1,400 aa, y cada monómero tiene una masa de alrededor de 180 kilodaltones (kDa)<sup>6,21</sup>. Esta organizada en dos subunidades, S1 y S2. La S1, a su vez, tiene dos dominios: un dominio N-terminal (NTD, por sus siglas en inglés) y un dominio de unión al receptor (RBD, por sus siglas en inglés), que es responsable de la interacción con el receptor de ACE2 para permitir la entrada de partículas virales a la célula. La subunidad S2 interviene en la fusión de las membranas viral y plasmática. Contiene un péptido de fusión (FP, por sus siglas en inglés), la hélice central (CH, por sus siglas en inglés) y el dominio de conexión (CD, por sus siglas en inglés). Los dominios adicionales dentro de la subunidad S2 incluyen los dominios de repetición de heptada 1 y 2 (HR1, HR2, por sus siglas en inglés) y los dominios transmembranales (TM, por sus siglas en inglés) que forman el tallo y la cola citoplasmática (CT, por sus siglas en inglés) (Figura 3). Entre ambas subunidades se encuentra un sitio de escisión de furina, que se

considera prepara a la proteína para la infección. Una segunda escisión proteolítica en el sitio S2 libera el FP, que penetra en la membrana de la célula huésped, preparándola para la fusión<sup>22</sup>.



**Figura 3. Estructura del monómero de la proteína S del virus SARS-CoV-2.** En color rojo se muestra la región de unión al receptor (RBD, 336-518), en azul oscuro el dominio N-terminal (NTD, 16-291), en magenta el sitio de escisión de la furina (S1/S2), en cian el péptido de fusión (FP, 817-834), en naranja la hélice central (CH, 987-1034), en verde el dominio de conexión (CD, 1080-1135), en café y amarillo los dominios de repetición de heptada 1 (HR1, 1047-1163) y 2 (HR2, 1163-1210) respectivamente, en azul claro el dominio transmembranal (TM, 1214-1234), y en morado la cola citoplasmática (CT, 1235-1273). (Obtenida de la colección de sistemas proteicos de SARS-CoV-2 de la plataforma web CHARMM-GUI usando como guía lo presentando por Woo y colaboradores y editada en el programa PyMol).

### **Proteína de Membrana (M)**

La glicoproteína M, de aproximadamente 25 kDa, es la proteína más abundante del virión, proporcionando una forma definida a la envoltura viral. Posee una longitud de aproximadamente 220-260 aa con un ectodominio N-terminal de longitud corta, integrado en la membrana del virus por medio de tres dominios transmembranales designados como TM1, TM2 y TM3; y un endodominio C-terminal. La glicoproteína M cumple funciones en la unión a la nucleocápside y actúa como un organizador del ensamblaje de los viriones dentro de la célula infectada; además, es necesaria para la retención de la glicoproteína S en el retículo endoplásmico-Golgi (ERGIC, por sus siglas en inglés) y su incorporación a nuevos viriones<sup>21,23,24,25</sup>.

### **Proteína de Envoltura (E)**

La proteína E, o de envoltura, es la más pequeña de las principales proteínas estructurales, es un pentámero con 76 aa de longitud y con monómeros de masa inferior a 8 kDa<sup>21,23</sup>. Su estructura comprende tres dominios: un ectodominio hidrófilo N-terminal que consta de 7 a 12 aa, un dominio transmembranal hidrófobo y un endodominio hidrófilo C-terminal largo<sup>25</sup>. Desempeña funciones en la patogenia, el ensamblaje, la liberación del virus y, adicionalmente, se ha reportado que actúa como viroporina (canal iónico)<sup>21,23,24</sup>.

### **Proteína Nucleocápside (N)**

La proteína N, de alrededor de 46 kDa, está conformada por dos dominios conservados: el dominio N-terminal y C-terminal<sup>21,23,25</sup>. Entre sus funciones, desempeña un papel en la formación de complejos con el genoma viral, facilita la interacción de la proteína M necesaria durante el ensamblaje del virión y mejora la eficiencia de transcripción del virus<sup>21,23,24</sup>.

## **Proteínas no estructurales y accesorias**

Como se mencionó anteriormente, el genoma de SARS-CoV-2, contiene 16 nsps, nsp1 a nsp11 y nsp12 a nsp16; estas cumplen un rol importante en la replicación del virus dentro de las células huésped y algunas de ellas son el objetivo de diversos fármacos en desarrollo<sup>21</sup>.

Las proteínas accesorias del SARS-CoV-2 son expresadas por los genes ORF1ab, ORF3a, ORF3b, ORF6, ORF7a, ORF7b, ORF8, ORF9a, ORF9b y ORF10. Varias de estas proteínas tienen funciones aún desconocidas, algunos estudios indican que podrían jugar un papel específico en la replicación viral<sup>21</sup>, mientras que otros mencionan que participan en la patogenicidad viral al modular la vía de señalización del interferón<sup>26</sup>. Las características que se han descrito de algunas de estas proteínas se detallan a continuación:

### **ORF1ab**

El marco de lectura abierto 1 (ORF1) funciona como replicasa, replicasa/transcriptasa o poliproteínas polimerasas, y se traduce en las poliproteínas ORF1a (486 kDa) y ORF1b (306 kDa). Las nsps participan en el ciclo biológico de SARS-CoV-2, siendo nsp3 y nsp5 responsables de la escisión de las poliproteínas 1a y 1b para la posterior formación del complejo replicasa-transcriptasa, y favoreciendo el avance de la infección viral<sup>21</sup>.

### **ORF3a**

La proteína ORF3a tiene una estructura de tres hélices transmembranales (TM1, TM2 y TM3) seguidas de un dominio citosólico de dos láminas  $\beta$  antiparalelas que forman un sándwich  $\beta$ . Una reciente estructura de ORF3a mostró que forma un dímero y las seis hélices transmembranales del dímero forman un canal iónico con residuos polares capaces de conducir cationes, por lo que se le consideró como la primera proteína de la familia de proteínas de coronavirus de las viroporinas<sup>25</sup>.

ORF3a está implicada en la inducción de apoptosis, patogenicidad, la formación de viroporinas sensibles al potasio y puede modular la liberación de virus<sup>24,25</sup>.

## **ORF6**

Se caracteriza por tener entre 42 y 63 aa de longitud<sup>27</sup>, cumple con funciones como la retención de factores de importación en la membrana del ERGIC conduciendo a una pérdida de transporte al núcleo. De este modo previene la translocación nuclear de STAT1 en respuesta a la señalización de interferón, bloqueando así la expresión de genes estimulados por interferón que muestran múltiples actividades antivirales<sup>24</sup>.

## **ORF7a y 7b**

El análisis de la secuencia de ORF7a indica que éste codifica una proteína de 121 residuos con un péptido señal N-terminal, un ectodominio, una región transmembranal y una cola C-terminal corta. La estructura muestra un pliegue sándwich  $\beta$  compacto de siete hebras<sup>25</sup>. Se ha descrito esta proteína en cultivos celulares como prescindible para la replicación del virus<sup>24</sup>, mientras que la proteína ORF7b se caracteriza por ser una proteína accesoria viral codificada en el ARN subgenómico<sup>27</sup>.

## **ORF8**

ORF8 es una proteína accesoria de 121 residuos que tiene una secuencia señal N-terminal y, a diferencia de las otras proteínas accesorias descritas anteriormente, carece del dominio transmembranal C-terminal. Las estructuras cristalizadas muestran que forma un dímero con un núcleo monomérico de 60 residuos. Cada monómero tiene ocho láminas  $\beta$  antiparalelas<sup>25</sup>.

Entre sus funciones, puede desempeñar un papel en la interacción virus-huésped, también se ha propuesto que cuando ORF8 se sobre-expresa



exógenamente en las células, interrumpe la señalización de IFN-I, asimismo regula a la baja la expresión de MHC-I en las células. No obstante, la función biológica exacta de ORF8 sigue sin estar clara y requiere de más investigación<sup>24,25</sup>.

## **ORF10**

La proteína ORF10 se encuentra río arriba de la región 3' no traducida (3'-UTR), que codifica una proteína de 38 aa de longitud<sup>27</sup>. Los estudios enfocados en esta proteína continúan, ya que aún no queda claro si esta región se traduce en una proteína funcional<sup>24</sup>.

### **1.3. Ciclo biológico de SARS-CoV-2**

El SARS-CoV-2 y los virus relacionados han evolucionado para aprovechar receptores en su propio beneficio al utilizarlos para la entrada y su propagación en las células humanas<sup>20</sup>. La glicoproteína S de SARS-CoV-2 se une al receptor celular ACE2. Los receptores ACE2 están presentes en las células caliciformes (secretoras) de las células ciliadas en la nariz, la parte posterior de la garganta, los pulmones, el intestino, los músculos de corazón y riñones, lo que facilita la ruta de transmisión y lo que podría explicar el fácil y rápido contagio del nuevo coronavirus<sup>28</sup>. Como se ha mencionado, la glicoproteína S utiliza sus dos subunidades funcionales para lograr la entrada del virus, la subunidad S1 es responsable de la unión con el receptor ACE2 de la célula hospedera, mientras que S2 facilita la fusión del virus con las membranas celulares<sup>18,22</sup>. Cuando la glicoproteína S de SARS-CoV-2 se une al receptor ACE2, el complejo resultante es procesado proteolíticamente por la proteasa transmembranal de serina tipo 2 (TMPRSS2, por sus siglas en inglés), lo que conduce a la escisión de ACE2 y a la activación de la glicoproteína S, comenzando el proceso de unión y fusión del virus con la membrana celular y finalizando con la entrada del virus a la célula hospedera<sup>18</sup>.

Una vez que se completa la unión virus-membrana celular, inicia la fusión del virus con ésta. Al finalizar la fusión, el siguiente paso en el ciclo de vida es la

liberación en el citoplasma del ARN genómico del virión y permitir la formación de las poliproteínas (pp) 1a y 1ab, así como la síntesis de ARN subgenómico y la replicación del genoma viral. Posteriormente, las proteínas estructurales S, E y M se traducen e insertan en el retículo endoplásmico. Estas proteínas se mueven a lo largo de la vía secretora hacia el ERGIC. Allí, los genomas virales son encapsulados por la proteína N y puestos dentro de las membranas del ERGIC, que contienen proteínas estructurales virales, formando así viriones maduros<sup>18,20</sup>.

Los viriones, una vez estructurados, se transportan a la superficie celular en vesículas, liberándose por exocitosis y proceden a infectar nuevas células. La infección conduce finalmente a la muerte de la célula, posiblemente por apoptosis, permitiendo que el virus se propague dentro de un organismo infectado sin ser detectado o neutralizado por anticuerpos específicos contra el virus<sup>18,20</sup>.

Hoy en día, seguimos en la batalla para detener la pandemia provocada por el SARS-CoV-2, debido a la aparición de cinco nuevas variantes de preocupación (Alfa, Beta, Gama, Delta y Omicron) y siete variantes de interés (Epsilon, Eta, Iota, Kappa, Lambda, Theta y Zeta)<sup>29</sup> y es probable que este número aumente en el futuro mientras el virus siga circulando activamente. No obstante, la vacunación sigue siendo una de las herramientas más importantes para controlar la pandemia. A continuación, se detallan los aspectos que están directamente implicados en la respuesta inmunitaria frente a las vacunas y al SARS-CoV-2.

## **2. Respuesta inmunológica frente a SARS-CoV-2**

La prevención de las infecciones mediante la vacunación es una estrategia de ayuda al sistema inmunológico, esto a través de la exposición previa a antígenos (fragmentos) del agente infeccioso o bien al agente infeccioso completo (muerto, atenuado o inactivado) que estimulan la respuesta inmune previamente a la llegada del agente infeccioso completo. El sistema inmunológico responde de dos maneras, la respuesta inmune innata y la respuesta inmune adaptativa<sup>30</sup>. La primera de ellas brinda una respuesta temprana generalizada que conduce a la inflamación; mientras que la segunda genera una respuesta específica frente a los antígenos<sup>7</sup> y genera memoria inmunológica<sup>31</sup>.

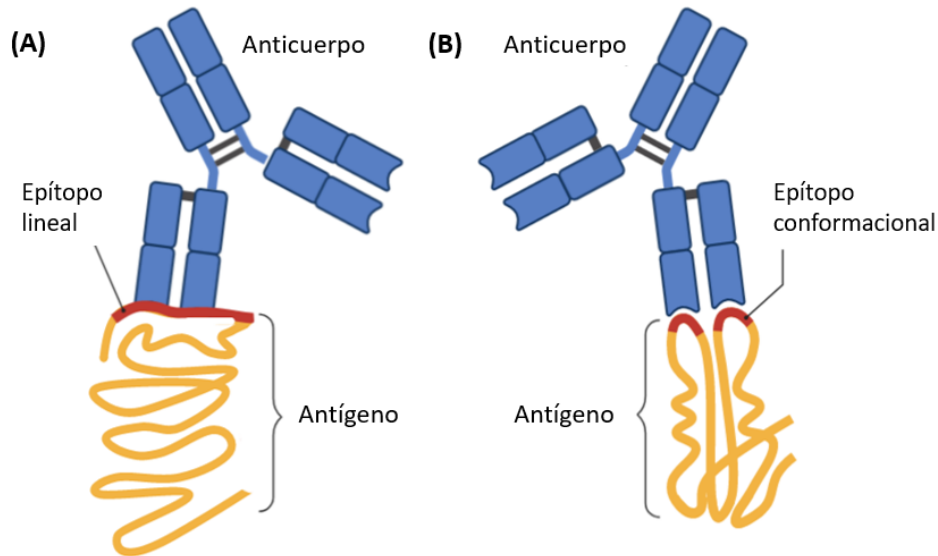
Las vacunas funcionan estimulando la respuesta inmune innata, activando a las células presentadoras de antígeno (CPAs): macrófagos, células dendríticas y células B; desencadenando una respuesta específica hacia un epítipo determinado en un antígeno. Posteriormente, se activan los mecanismos de inmunidad adaptativa, cuya respuesta involucra dos tipos de células, la célula B y la célula T<sup>7,30</sup>. La célula T induce mecanismos efectores que eliminan a las células infectadas y los patógenos, mientras que las células B producen anticuerpos específicos contra los mismos. Algunas de las células T y células B forman células de memoria que se almacenan en los ganglios linfáticos, en caso de que la persona esté infectada con el mismo patógeno, las células de memoria producen anticuerpos contra la infección eliminando el agente infeccioso<sup>7</sup>.

Para la eliminación de las células infectadas por el virus SARS-CoV-2 son esenciales ambas respuestas. Por un lado, las respuestas inmunitarias adaptativas mediadas por células T CD4+ y CD8+ son esenciales para promover respuestas eficientes de células B y anticuerpos, así como para la eliminación de células infectadas por el virus SARS-CoV-2, mientras que las respuestas de células B y anticuerpos son importantes para generar anticuerpos neutralizantes. Se ha descrito que la producción de anticuerpos contra la infección por SARS-CoV-2 comienza con la inmunoglobulina M (IgM) donde hay un aumento en el noveno día, mientras que la presencia de la inmunoglobulina G (IgG) y de la inmunoglobulina A (IgA) comienza después de los siete días de infección<sup>32</sup>. Además, el período de incubación de este virus se asocia con la supresión de la respuesta inmune innata y la activación retardada de las células T (CD8+), siendo que este mismo mecanismo se ha descrito para el SARS y MERS. Al comparar los casos leves con los casos graves, se ha observado que en los primeros se tiene una mayor cantidad de células T de memoria CD8+ en el tracto respiratorio<sup>33</sup>, mientras que la infección grave por SARS CoV-2 se caracteriza por el proceso de linfopenia con recuentos muy bajos de células T CD4+, células T CD8+, células B, células NK, así como una mayor proporción de neutrófilos en comparación con los linfocitos<sup>34</sup>.

## 2.1. Determinante antigénico o epítopo

Los determinantes antigénicos o epítomos son las regiones de un antígeno o inmunógeno que interaccionan con el sitio activo, o paratopo, reconocidas como extrañas y provocando una respuesta inmune identificada por las células B y T. Por lo anterior, no todo el inmunógeno interacciona con un anticuerpo, sino sólo el epítopo. Por definición, para ser antígeno o inmunógeno una molécula deberá tener, al menos, un epítopo; entre más compleja sea la molécula, tendrá más epítomos<sup>7</sup>.

El grupo de Sela estudió a los epítomos y los clasificó de la siguiente manera: los epítomos *secuenciales*, *continuos* o *lineales*, los cuales dependen del orden en el que se acomodan los monómeros de la macromolécula; en el caso de las proteínas, es la estructura primaria, y la diferencia de un solo aminoácido puede ser suficiente para inducir una respuesta inmunológica (Figura 4-A). Los epítomos *conformacionales* o *discontinuos* dependen de la conformación tridimensional que acerca dos segmentos lejanos de una molécula de manera secuencial; la forma en que se genera en el espacio debido a esta cercanía es reconocida por el sitio activo del receptor de la célula B (BCR) y de los anticuerpos. Estos epítomos son característicos de las proteínas y se dan por estructuras terciarias o cuaternarias (Figura 4-B)<sup>7,35</sup>.



**Figura 4. Representación esquemática de los dos tipos de epítopos presentes en los antígenos.** (A) Un epítipo lineal (en rojo) presenta estructura primaria dada por la secuencia de aminoácidos, la cual se reconoce por un anticuerpo específico (en azul). (B) Un epítipo conformacional (en rojo) se genera por estructuras secundarias, terciarias o cuaternarias.

## 2.2. Reactividad cruzada

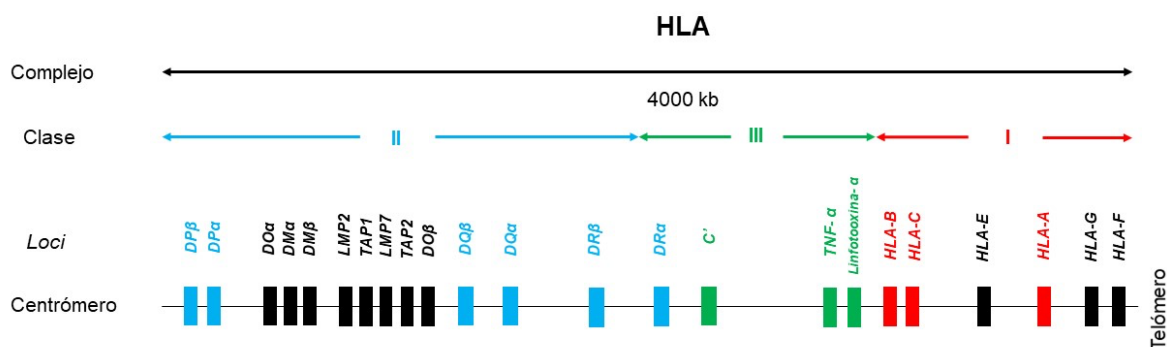
Cuando algunos epítopos de un antígeno *X* son compartidos por otro antígeno *Y*, una proporción de los anticuerpos dirigidos contra *X* también reaccionarán contra *Y*, a esto se le denomina reactividad cruzada<sup>7</sup>.

Diversos estudios sobre el análisis filogenético de la secuencia de la glicoproteína S de SARS-CoV-2 comparada con otros coronavirus (HCoV-NL63, HCoV-229E, HCoV-OC43, HCoV-HKU1, SARS-CoV y MERS-CoV) revelan tener la ascendencia más común con SARS-CoV y MERS-CoV. De manera similar, otro análisis filogenético realizado de diferentes variantes de SARS-CoV-2 aisladas de diferentes países, mostró que, en efecto, todas las glicoproteínas de las diferentes variantes de SARS-CoV-2, están estrechamente relacionadas entre sí, con un porcentaje de similitud general del 75%, lo que sugiere que una vacuna diseñada contra una variante también sería eficaz contra todas las demás variantes del virus aisladas de diferentes países gracias a la reactividad cruzada<sup>35</sup>, siempre y cuando se elija el epítipo más conservado en las diferentes variantes.

### 2.3. Complejo principal de histocompatibilidad (MHC)

El sistema inmunológico muestra una gran versatilidad para reconocer y responder a una gran cantidad de antígenos y para eliminarlos, pero también evita reaccionar con constituyentes propios. Esto se realiza por la colaboración de las células B y T, sin embargo, lo hacen de manera distinta. Mientras que las células B reaccionan con el antígeno directamente, las células T reconocen el antígeno procesado y presentado por una CPA en asociación con el Complejo Principal de Histocompatibilidad (MHC, por sus siglas en inglés)<sup>36</sup>.

El MHC es un conjunto de genes dispuesto dentro de un tramo continuo largo de ADN en el cromosoma 6 de los seres humanos, el cual es denominado complejo de antígeno leucocitario humano (HLA, por sus siglas en inglés). El *locus* del MHC codifica para dos clases principales de moléculas: las moléculas de clase I y las moléculas de clase II, ambas son glicoproteínas unidas a membrana que están relacionadas tanto en estructura como función (Figura 5)<sup>7</sup>.



**Figura 5. Mapa de los *loci* de MHC del cromosoma 6 en el ser humano.** Los genes que codifican para MHC-I están coloreados de rojo, los genes que codifican para MHC-II están coloreados de azul (Adaptada y traducida de Kuby, 2013).

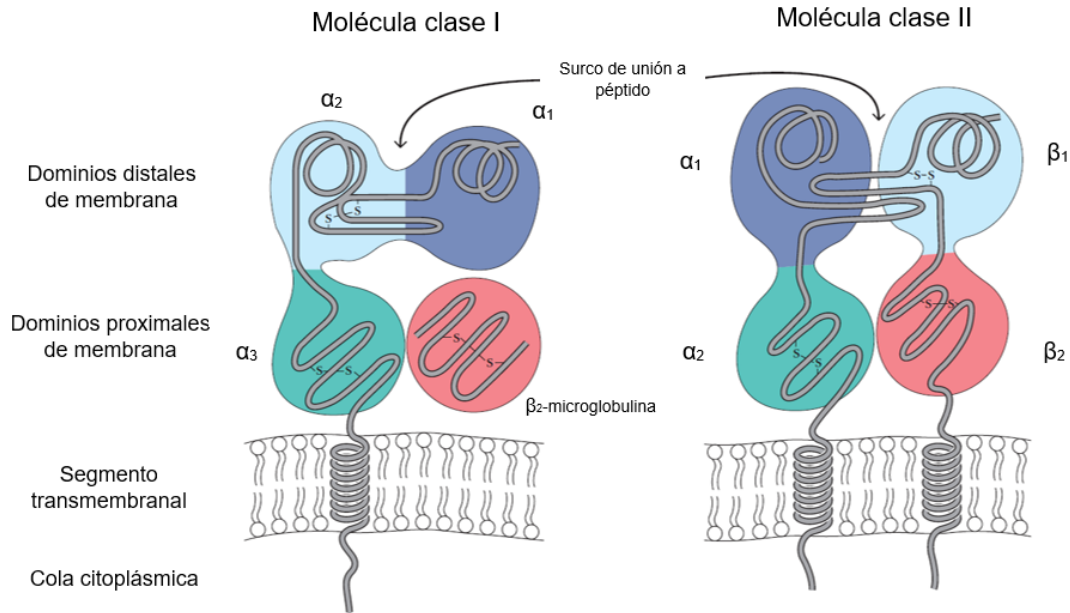
Las moléculas de clase I (MHC-I), formadas por una cadena  $\alpha$  de 45 kDa y una molécula de  $\beta_2$  microglobulina de 12 kDa (Figura 6), están presentes sobre la superficie de casi todas las células nucleadas y se especializan en presentar

antígenos que se originan a partir del citosol, como proteínas virales. Estas son presentadas en péptidos de 8 a 11 aa de longitud, por medio del surco de unión a péptido, a células T CD8+ que reconocen y matan las células que expresan esos antígenos intracelulares. Las moléculas de cadena  $\alpha$  están codificadas por los *loci* A, B y C en seres humanos. En conjunto, estas se denominan moléculas clase I clásicas<sup>7,37</sup>.

En contraste, las moléculas de clase II (MHC-II) contienen dos cadenas polipeptídicas diferentes, una cadena  $\alpha$  de 33 kDa, y una cadena  $\beta$  de 28 kDa (Figura 6), las cuales son expresadas de manera predominante sobre CPAs<sup>7,37</sup>. Las CPAs se definen como aquellas células que presentan antígenos en contexto de moléculas del MHC a linfocitos T con receptores específicos para el antígeno<sup>38</sup>.

Una vez expresada sobre la superficie celular, la molécula de MHC-II presenta péptidos antigénicos más largos de 13 a 17 aa de longitud a células T CD4+, que entonces quedan activadas y proceden a destruir invasores extracelulares. Las moléculas de MHC-II son codificadas por las regiones DP, DQ y DR en humanos y se denominan moléculas clase II no clásicas<sup>7,37</sup>.

Las células T CD4+ también son responsables de la activación de células B, las cuales reconocen epítomos tanto lineales como conformacionales ya sea por el BCR o por un anticuerpo<sup>7,37</sup>.



**Figura 6. Diagrama de las moléculas de MHC clase I y clase II.** El surco de unión a péptido está formado por los dominios distales (en azul) de membrana en moléculas tanto clase I como clase II. Se muestran los dominios proximales de membrana (en verde y rojo), los segmentos transmembranales y las colas citoplásmicas. (Adaptada y traducida de Kuby, 2013).

Cualquier individuo sólo expresa un pequeño número de MHC, hasta 8 moléculas de clase I y 13 o más moléculas clase II diferentes. Aun así, este número limitado de moléculas de MHC debe ser capaz de presentar una enorme gama de diferentes péptidos antigénicos a células T, lo que permite que el sistema inmunitario muestre respuesta de manera específica a una amplia variedad de desafíos antigénicos. De este modo, la unión a péptido por moléculas de clases I y II no muestra la especificidad fina característica de la unión a antígeno por anticuerpos y receptores de la célula T. En lugar de eso, una molécula de MHC dada puede unirse a varios péptidos diferentes, y algunos péptidos pueden unirse a varias moléculas de MHC distintas. Debido a esta especificidad amplia, la unión entre un péptido y una molécula de MHC a menudo se denomina “promiscua”. Esta promiscuidad de la unión a péptido permite que muchos péptidos diferentes coincidan con el surco de unión de MHC, y que en ocasiones suceda intercambio de péptidos, a diferencia de las interacciones cognadas de alta afinidad,



relativamente estables, de anticuerpos y receptores de célula T con sus ligandos específicos<sup>7,37</sup>.

En resumen, tanto la reactividad cruzada como la promiscuidad del MHC favorecen que la inmunidad generada contra un epítipo de la variante X puede emplearse para atacar otras variantes (Y, Z, etcétera), por lo que una clave importante y también uno de los primeros pasos en la generación de las vacunas es identificar el “epítipo ideal”, el cual debe ser una secuencia altamente conservada y que genere la mayor inmunogenicidad posible. La búsqueda del “epítipo ideal” ha aumentado el desarrollo de programas o *softwares* para la predicción de epítipos, algunos de los cuales se discuten en las siguientes secciones.

### **3. Predicción bioinformática de epítipos de SARS-CoV-2**

Poco después de la aparición de las primeras secuencias genéticas del SARS-CoV-2, comenzaron a emplearse métodos *in silico* para predecir y recomendar epítipos de células B y T como objetivos potenciales para el diseño y desarrollo de vacunas contra el SARS-CoV-2. Muchas de estas predicciones han sido útiles para informar los estudios experimentales que van surgiendo actualmente y que favorecen la comprensión de las respuestas inmunitarias provocadas naturalmente en pacientes convalecientes de COVID-19<sup>39</sup>.

Las vacunas desarrolladas utilizan enfoques convencionales que se basan en el uso de virus atenuados o inactivados, sin embargo, estos enfoques no han producido vacunas exitosas contra otros virus letales como el virus de la inmunodeficiencia humana (VIH), el virus de la hepatitis C y el virus del dengue. En comparación, las tecnologías de vacunas modernas, por ejemplo, las que se basan en péptidos, vacunas de ADN o ARN y las vacunas de subunidades proteicas, ofrecen alternativas prometedoras. De echo, los estudios clínicos en fase II y III para SARS-CoV-2 se basan en enfoques vacunales modernos. Estos enfoques introducen partes específicas del virus o sus genes en el cuerpo para producir una respuesta inmunitaria específica<sup>39</sup>.

El desafío común que enfrentan los enfoques vacunales modernos es identificar péptidos específicos de SARS-CoV-2 que sean capaces de inducir una respuesta inmune de células B y T protectora. Esta es una tarea imprescindible en el desarrollo de vacunas, ya que proporciona las recomendaciones básicas para todos los esfuerzos posteriores que se llevarán a cabo en la síntesis de vacunas, las pruebas de laboratorio y los ensayos clínicos<sup>39</sup>.

La identificación experimental de péptidos del SARS-CoV-2 que provoquen respuestas importantes de células B y T es difícil, debido a la gran cantidad de opciones posibles para probar, sin embargo, estos problemas se pueden resolver por medio de la identificación de epítomos de células B y T que aprovechen estos datos a través de métodos *in silico*. Varios autores han informado el desarrollo de vacunas candidatas basadas en epítomos mediante el uso de enfoques bioinformáticos. Los cuales han predicho péptidos con potencial para interactuar con MHC-I y MHC-II, mostrando respuestas inmunogénicas, no toxicidad y no alergenicidad. Además, del potencial para inducir la liberación de IFN $\gamma$  de los epítomos prometedores<sup>40</sup>. En otro estudio, se seleccionaron péptidos que contenían epítomos lineales de células B y epítomos de células T predichos por métodos bioinformáticos para el desarrollo de vacunas basadas en “multiepítomos”. Los resultados arrojaron que estos péptidos tenían una afinidad de unión fuerte a una serie de moléculas MHC, por lo que es probable que los epítomos tengan efectos de protección cruzada contra el SARS-CoV-2<sup>41</sup>.

En comparación con el desarrollo tradicional de vacunas, la predicción de epítomos mediante análisis bioinformático permite que el diseño de vacunas sea sencillo y rápido, con el fin de guiar a los investigadores para que tomen una decisión informada sobre la selección de vacunas candidatas, que deberá validarse sucesivamente en modelos *in vitro* e *in vivo*. A continuación se describen los métodos de predicción utilizados en este estudio.

### **3.1. Métodos de predicción de epítomos**

En la actualidad, se han desarrollado nuevas estrategias computacionales (bases de datos, métodos predictivos, softwares, etc.) que permiten predecir

epítomos potenciales capaces de desencadenar una respuesta humoral y celular. Por lo anterior, la bioinformática abre las posibilidades para optimizar el diseño de vacunas *in silico*, las cuales parecen ser prometedoras, especialmente para los patógenos emergentes<sup>21</sup>.

La predicción de epítomos se remonta a 1981, cuando Hopp y Woods desarrollaron el primer método de predicción de epítomos de células B. Desde entonces se han desarrollado o adaptado muchos más métodos a partir de otras herramientas computacionales<sup>42</sup>. Por ejemplo, BcePred es un servidor que predice epítomos de células B basados en las propiedades fisicoquímicas de los aminoácidos como su hidrofiliidad, flexibilidad, accesibilidad, polaridad, propensión antigénica e, incluso, una combinación de todas<sup>43</sup>. Otros son los servidores Pepitope y EpiSearch que se utilizan para mapear los péptidos en estructuras proteicas antigénicas en 3D con el fin de aclarar regiones específicas de interés de los epítomos<sup>44</sup>. Otra, es la base de datos SYFPEITHI con más de 7,000 péptidos con afinidad por MHC-I y MHC-II. Sus predicciones se basan sólo de informes publicados y en lo que respecta a los epítomos de células T, permite una selección de aquellos que probablemente se procesen naturalmente, teniendo en cuenta los aminoácidos en las posiciones de anclaje, así como otros aminoácidos frecuentes<sup>45</sup>. Por último, el servidor web IFNepitope (<http://crdd.osdd.net/raghava/ifnepitope/>), permite predecir y diseñar péptidos según su potencial para inducir interferón gamma (IFN $\gamma$ )<sup>40</sup>.

En este estudio se utilizó el recurso de análisis y bases de datos de epítomos inmunes (IEDB, por sus siglas en inglés) (<https://www.iedb.org/>), el cual es una plataforma que contiene epítomos tanto de células B como de células T para enfermedades infecciosas, alérgenos, enfermedades autoinmunes y trasplantes que son estudiados experimentalmente en humanos, primates, ratones y otras especies animales. Asimismo, cuenta con un conjunto de herramientas para predecir y analizar más de 260,000 epítomos, la mayoría de los cuales se extraen de publicaciones científicas y, a diferencia de SYFPEITHI, incluye una gran cantidad de datos sobre péptidos sintéticos<sup>46</sup>.

Las herramientas para predecir epítomos lineales de células B, la unión de MHC-I y MHC-II y la inmunogenicidad, utilizadas en este estudio se describen a continuación:

### **3.2. Predicción de epítomos de células B**

El IEDB alberga las herramientas BepiPred y BepiPred 2.0 (<http://tools.iedb.org/bcell/>) para la predicción de epítomos lineales de células B. El primero es un servidor de predicción, hecho mediante la combinación de las predicciones de un modelo de Markov oculto y la escala de propensión de Parker. Dicho servidor ha demostrado que funciona significativamente mejor en comparación con otros métodos disponibles en la página web, por esta razón BepiPred es el método predeterminado<sup>46,47</sup>. En comparación, el servidor web BepiPred 2.0 se basa en un algoritmo forestal aleatorio entrenado en epítomos anotados a partir de estructuras proteicas de anticuerpos y antígenos; superando a otras herramientas disponibles para la predicción de epítomos derivados de una gran colección de epítomos lineales descargados de la base de datos del IEDB<sup>48</sup>.

Los resultados que arrojan estos servidores contienen un gráfico de puntuación de las posiciones pronosticadas de los residuos de los epítomos basados en el valor umbral, este valor se basa en la sensibilidad y especificidad, para BepiPred un valor umbral de 0.35 corresponde a una sensibilidad de 0.49 y una especificidad de 0.75. El aumento del umbral reduce la sensibilidad y aumenta la especificidad, lo que reduce los falsos positivos. Debajo del gráfico se genera, también, una tabla que enumera los epítomos predichos, la secuencia, sus posiciones en la proteína y longitud. Finalmente, una segunda tabla enumera la puntuación predicha para el residuo en cada posición, y una columna de asignación que indica la posición del epítomo predicho señalada con una letra “E”<sup>46,47,49</sup>.

### **3.3. Predicción de epítomos de células T**

Las herramientas utilizadas en este estudio para la predicción de epítomos de células T, específicamente para la predicción de unión de MHC-I y MHC-II,

incluyendo TepiTool y la predicción de la inmunogenicidad de los epítomos, se detalla a continuación:

### 3.3.1. Predicción de epítomos de unión de MHC-I y MHC-II

Los métodos de predicción de unión de MHC facilitan la selección de epítomos potenciales. Dichos métodos utilizan datos experimentales de unión de péptidos para diferentes alelos del MHC para entrenar algoritmos de aprendizaje automático que, a su vez, pueden usarse para predecir la unión de cualquier péptido arbitrario. Los enfoques de aprendizaje automático pueden desarrollar una función que predice la unión por afinidad para una secuencia de péptidos determinada. De este modo, la afinidad entre un MHC y péptidos aislados es expresada como concentración  $IC_{50}$ . Un valor  $IC_{50}$  bajo, implica una alta afinidad de unión del péptido<sup>47</sup>. Los métodos de aprendizaje automático se capacitan periódicamente cuando hay suficiente cantidad de datos nuevos disponibles en el IEDB<sup>46</sup>.

La precisión de las herramientas para la clase I es significativamente alta, cubriendo 173 moléculas de MHC-I, que incluyen 119 alelos humanos para HLA-A, B y C<sup>46</sup>. Para realizar estas predicciones se ingresan las secuencias de las proteínas de interés en la página web de inicio de clase I (<http://tools.iedb.org/mhci/>), una vez ingresadas, se selecciona un método de una lista que incluye el recomendado por IEDB, consenso, netMHCpan, red neuronal artificial (ANN), método de matriz de puntuación (SMM), SMM con una matriz de covarianza de energía de unión de péptido (SMMPMBEC), Biblioteca combinatoria, PickPocket y netMHCcons (Figura 7).

El método netMHCpan genera predicciones cuantitativas de la afinidad de cualquier interacción péptido-MHC-I, del cual ya se ha demostrado con gran precisión la unión de péptidos a moléculas MHC-I en chimpancés y macacos<sup>50</sup>. A diferencia de este, PickPocket y netMHCcons, pueden predecir la unión del péptido también a las moléculas del MHC para las cuales los datos son limitados o ningún dato experimental de la unión del péptido está disponible<sup>1</sup>. El otro método es la red neuronal artificial (ANN) que se caracteriza por reconocer efectos correlacionados donde la afinidad de unión de un aa dado en una posición está influenciado por aa

en otras posiciones en el péptido. Dos aa adyacentes pueden, por ejemplo, competir por el espacio en la molécula de MHC. Por lo anterior, ANN es ideal para tener en cuenta dichas correlaciones, así como para predecir si un péptido se une o no a las moléculas de MHC<sup>52</sup>. Los métodos SMM y SMMPMBEC implican el reconocimiento y unión de péptidos inmunogénicos a moléculas de MHC. A partir de datos experimentales, se construye un modelo de la especificidad de secuencia de estos procesos. Su propósito es proporcionar un resumen de los resultados experimentales permitiendo una comprensión más profunda de los mecanismos involucrados en el reconocimiento de secuencias. Así como para predecir el resultado experimental de secuencias aún no probadas<sup>53</sup>. El otro método es la Biblioteca combinatoria, la cual es una herramienta útil para caracterizar la especificidad de unión de las moléculas de MHC-I y en comparación con otras metodologías, como la medición de las afinidades de péptidos individuales, el uso de bibliotecas combinatorias de exploración posicional proporciona una caracterización de referencia de la especificidad molecular del MHC que es rentable, cuantitativa e imparcial<sup>54</sup>.

El método recomendado por IEDB es el consenso, el cual se utilizó en este estudio y que tiene la ventaja de ser una combinación de tres métodos que proporcionan el mejor rendimiento (ANN, SMM y Biblioteca combinatoria). Posteriormente, se especifican los alelos y longitudes de los epítomos. Para los seres humanos, los alelos más frecuentes (97% de la población mundial) están disponibles para su selección de forma predeterminada y longitudes de epítomos de 9 y 10 aa, de esta manera la secuencia de proteínas se analiza en todos los péptidos posibles para la longitud especificada y se calcula la afinidad de unión para cada uno<sup>46</sup>.

La herramienta compara la afinidad predicha con un conjunto de péptidos de tamaño similar seleccionados al azar y asigna un rango de percentil. Un rango de percentil bajo corresponde a una afinidad de unión más alta<sup>46</sup>.

Los resultados se presentan en una tabla que incluye columnas para el alelo, las posiciones inicial y final del péptido, la secuencia del péptido, los métodos utilizados y el rango de percentil<sup>46</sup>.

Finalmente, en cuanto a la clase I existen varias estrategias para la selección de péptidos potenciales, incluidas la selección del porcentaje deseado de los péptidos ordenados por rango de percentiles, dando como resultado un número fijo de los mejores péptidos posibles. Otro enfoque, es elegir todos los péptidos con un rango percentil menor o igual a 10, incluido un valor  $IC_{50}$  menor de 500 nM, un umbral asociado con la inmunogenicidad<sup>46</sup>, este último criterio fue el empleado para este estudio.

Por otra parte, la precisión de las predicciones para la clase II ha mejorado significativamente en los últimos años, sin embargo, es más baja que para la clase I. La amplitud de los datos de alelos también es menos extensa, involucrando 24 alelos humanos. Para llevar a cabo estas predicciones en la página web IEDB (<http://tools.iedb.org/mhcii/>) se ingresan las secuencias de interés, posteriormente, se selecciona el método de predicción que incluyen los recomendados por IEDB, consenso, netMHCIIpan, SMM-align, Sturniolo y NN-align (Figura 7). Diversos métodos de predicción de secuencias de péptidos-MHC-II a la vez de predecir los núcleos péptidos de una base de datos de entrenamiento generan un alineamiento de esas mismas secuencias, pudiendo hacer notar la similitud entre secuencias que tiene el conjunto de datos. El método de alineamiento de secuencias por estabilización de matrices (SMM-align) o el alineamiento por redes neuronales (NN-align) generan este tipo de alineamientos. El método predeterminado es el 2.22 recomendado por IEDB. Cada método genera un rango percentil y una puntuación de afinidad de unión<sup>46</sup>.

Con esta herramienta es posible especificar la especie, el lugar (DR, DP y DQ, para los humanos) y los alelos, además, se pueden seleccionar los alelos más comunes de un conjunto de 27 alelos que cubren el 94% de la población mundial. La herramienta analiza las secuencias de proteínas y predice la afinidad de unión para cada péptido. Luego compara la afinidad predicha para cada péptido con la de un gran conjunto de péptidos seleccionados al azar para determinar su rango percentil. Como se mencionó anteriormente, cuánto más bajo sea el rango de percentil, mayor será la afinidad de unión<sup>46</sup>.

Los resultados se muestran en una tabla enumerando el alelo en la primera columna, seguido de las posiciones de la secuencia inicial y final, la secuencia de péptidos, el método utilizado y el rango percentil<sup>46</sup>.

### 3.3.2. TepiTool

*TepiTool* (<http://tools.iedb.org/tepitool/>) es una herramienta para la predicción de unión de MHC-I y MHC-II que proporciona valores predeterminados recomendados por IEDB para la selección automática de péptidos óptimos. Dicha herramienta proporciona una guía a través de seis pasos para especificar las secuencias de proteínas, especies hospederas y clase de MHC, y alelos. Para los alelos de clase I es posible seleccionar de una lista, todos los alelos disponibles, incluso, los de mayor frecuencia (frecuencia superior al 1% en la población global). También es posible elegir el número de péptidos, así como las longitudes de estos<sup>46</sup>.

Posteriormente, se selecciona el método de predicción que incluyen el recomendado por IEDB, consenso, SMM o ANN, y los criterios de salida (rango percentil predicho, umbrales de unión específico de IC<sub>50</sub> o MHC pronosticados)<sup>46</sup>.

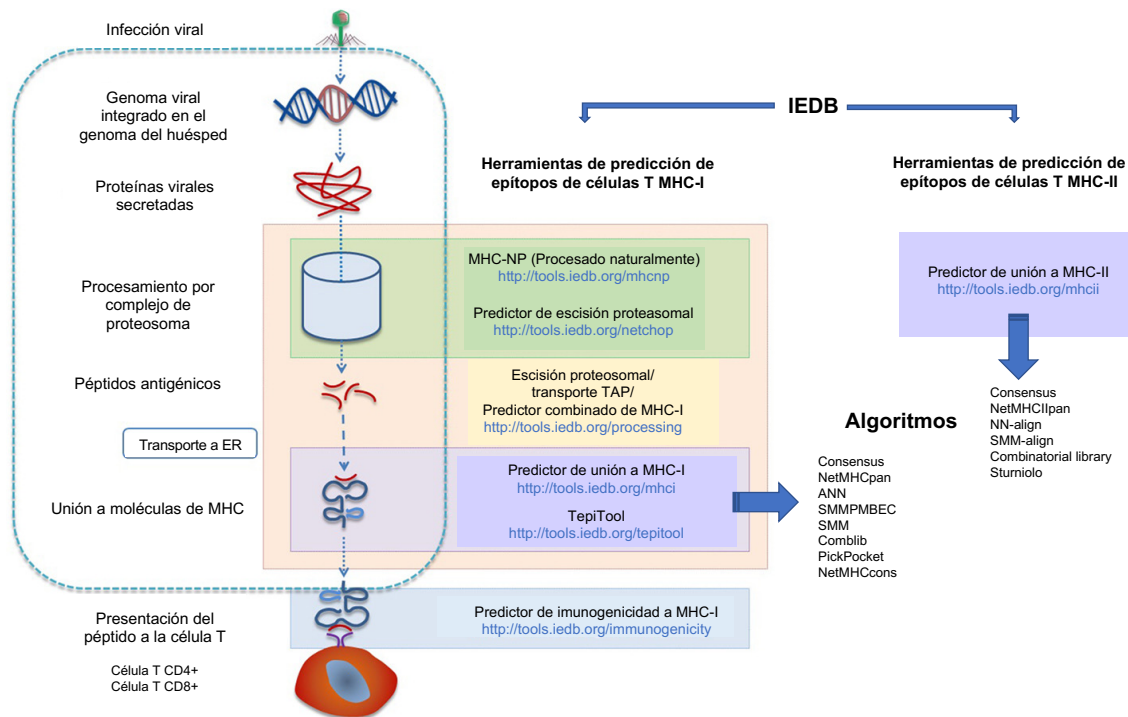
Los resultados se muestran en una tabla que contiene las posiciones inicial y final del péptido, la secuencia del péptido, el rango percentil y el nombre del alelo<sup>46</sup>.

Las predicciones de unión MHC-II se realizan de manera similar, con excepción de la selección del número de péptidos y de la elección de alelos donde es posible seleccionarlos de un panel de los 27 alelos más frecuentes<sup>46</sup>.

### 3.3.3. Inmunogenicidad

Esta herramienta (<http://tools.iedb.org/immunogenicity/>) predice la capacidad relativa de un conjunto determinado de péptidos que se unen a MHC-I para ser reconocidos por una célula T, de tal manera que los clasifica en dos categorías: epítomos y no epítomos. Dicha herramienta se basa en un análisis de la composición de aa del péptido en posiciones de no anclaje, donde es probable que las cadenas laterales de los aa estén en contacto con el TCR<sup>46,47</sup>.





**Figura 7. Diagrama de flujo de las herramientas de predicción disponibles en el recurso IEDB.** Del lado izquierdo se muestran las herramientas de predicción con respecto a las diferentes etapas del procesamiento del antígeno MHC-I y sus algoritmos antes descritos, del lado derecho se muestran las herramientas de predicción de epítopos de células T MHC-II y sus algoritmos. (Adaptada y traducida de Fleri *et al.*, 2017).

#### 4. Alineación de secuencias

La aparición de muchos genomas recién ensamblados permite abordar varias preguntas relacionadas con la evolución. Por ejemplo, la genómica comparada es un campo de la biología en la que se comparan las secuencias del genoma de diferentes especies para comprender cómo se relacionan los organismos entre sí a nivel genético. La detección de eventos de mutaciones puntuales es el objetivo de los estudios de la *alineación de secuencias*<sup>55</sup>.

En general, esta alineación permite conocer la historia evolutiva de los genes y los organismos de donde provienen estas secuencias, analizar filogenéticamente genes o proteínas homólogos, proporcionar información sobre la función, estructura, plegamiento y el sitio de unión de una proteína, así como de las posiciones de

nucleótidos que se han mantenido invariantes o donde han ocurrido cambios durante su historia evolutiva<sup>55,56</sup>.

Para detectar estos eventos existen bases de datos especializadas en secuencias genéticas y de proteínas a las que se puede acceder en línea. En este estudio se utilizaron dos programas para realizar alineamientos de las secuencias de las proteínas de SARS-CoV-2 y otros coronavirus humanos, que se describen a continuación:

#### **4.1. Clustal Omega**

Las secuencias de genes evolucionan bajo procesos que incluyen eventos como sustituciones, inserciones y eliminaciones que pueden cambiar la longitud de las secuencias y deben tenerse en cuenta para la genómica comparada. Estos eventos tienen el efecto de difuminar qué partes de secuencias de varios organismos están relacionados entre sí. Dos o más caracteres en una secuencia se denominan homólogas si ambas se derivan de un carácter común en un ancestro, y esta relación se llama homología<sup>55</sup>.

Clustal Omega se utiliza para crear múltiples alineaciones de secuencias (MSA, por sus siglas en inglés) (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Este es un procedimiento para alinear rápidamente y de forma precisa tres o más secuencias biológicas de longitud similar. A partir de la salida de las aplicaciones de MSA, se puede inferir la homología y la relación evolutiva entre las secuencias estudiadas. La velocidad proviene del uso del algoritmo mBed, que calcula árboles guía en tiempo y memoria  $O(N \log N)$  y la precisión proviene del uso de Hhalgn, un alineador sofisticado de modelo de Markov oculto (HMM) que alinea pares de HMM<sup>55,56</sup>.

#### **4.2. EMBOSS Needle**

EMBOSS Needle ([https://www.ebi.ac.uk/Tools/psa/emboss\\_needle/](https://www.ebi.ac.uk/Tools/psa/emboss_needle/)) es un programa que utiliza el algoritmo de alineación global de Needleman-Wunsch para encontrar la alineación óptima de dos secuencias al considerar su longitud

completa, explorando todas las alineaciones posibles y eligiendo la mejor. Lo hace leyendo una matriz de puntuación que contiene valores para cada posible residuo o coincidencia de nucleótidos. Needle encuentra una alineación con la puntuación máxima posible, donde la puntuación de dicha alineación es igual a la suma de las coincidencias tomadas de la matriz de puntuación. Asimismo, proporciona un porcentaje de similitud e identidad entre las secuencias<sup>57</sup>, el cual se utilizó como filtro final para la selección de los epítomos potenciales en este estudio.

## PLANTEAMIENTO DEL PROBLEMA

Desde finales de diciembre de 2019, el mundo se ve afectado por una situación de emergencia de salud pública debido a la pandemia viral causada por la enfermedad COVID-19. Ante esta situación se están desarrollando más de 200 vacunas candidatas que se encuentran en diferentes fases de desarrollo preclínico y clínico, y algunas de ellas ya han sido aprobadas para su uso de emergencia, mientras, otras ya cuentan con licencia. La mayoría de los esfuerzos se han centrado en el uso de la glicoproteína S de SARS-CoV-2 ya que, además, de ser la encargada de mediar la entrada del virus a la célula huésped y la proteína de superficie común a todos los miembros de la familia viral, se han reportado niveles de anticuerpos neutralizantes significativos contra S en pacientes recuperados, lo que puede ser clave importante de protección contra la infección. Sin embargo, también ya se han documentado diversas variantes del virus con mutaciones claves en esta proteína que afectan negativamente la capacidad neutralizante de los anticuerpos, lo que podría sugerir que en el futuro sean capaces de evadir la respuesta inmune de forma mayormente significativa.

Asimismo, se ha propuesto que los anticuerpos neutralizantes producidos por células B es uno de los correlatos de protección después de la vacunación o infección. A pesar de esto, se tiene la teoría de que las células T también tienen un papel importante en la protección, ya que existe evidencia experimental que sugiere que las respuestas de células T de reactividad cruzada entre coronavirus están presentes en la población, sin embargo, se desconoce cuáles son los posibles antígenos compartidos entre diferentes coronavirus que podrían contribuir a la protección.

Por lo anterior, consideramos que es de fundamental importancia analizar la inmunogenicidad de las proteínas que conforman el virus SARS-CoV-2, para conocer los posibles blancos a las respuestas cruzadas de células B y T en coronavirus relevantes para la salud humana.

## JUSTIFICACIÓN

Los recientes avances en biología computacional abren las posibilidades para diseñar vacunas efectivas *in silico*, esta es una de las herramientas que permitirá a la humanidad controlar la propagación de SARS-CoV-2 y otros coronavirus que circulan actualmente entre la población humana, o incluso aquellos que podrían surgir en el futuro de otras especies animales.

El desarrollo de vacunas basadas en péptidos, que comprenden fragmentos de péptidos inmunogénicos cortos, ofrece ventajas prospectivas sobre el enfoque de la proteína completa debido a que son capaces de provocar respuestas inmunes fuertes y dirigidas, mediadas tanto por anticuerpos como por células T.

Además, se tiene la teoría de que las respuestas de las células T juegan un papel clave en la inmunidad contra el SARS-CoV-2, debido a que se ha observado que existe una respuesta inmune cruzada entre otros coronavirus estacionales y el SARS-CoV-2, sin embargo, no se han definido los epítomos con potencial a generar esta respuesta cruzada.

A través del análisis bioinformático es posible construir un repertorio de epítomos potenciales de células B y T que se encuentren en las distintas proteínas que conforman al SARS-CoV-2, e identificar aquellos que se encuentran conservados en otros coronavirus relevantes para la salud humana. Dicho conocimiento es de gran importancia para el desarrollo racional de una vacuna que potencialmente proporcione protección “universal” contra este, e incluso, otros coronavirus humanos.

## **PREGUNTA DE INVESTIGACIÓN**

¿Cuáles son los epítomos potenciales de células B y T que pueden generar inmunidad cruzada contra SARS-CoV-2 y otros coronavirus humanos?

## **HIPÓTESIS**

A través de métodos bioinformáticos se encontrarán epítomos de células B y T de las proteínas S, M, N, E, ORF1ab, ORF3a, ORF6, ORF7, ORF7b, ORF8 y ORF10 del virus SARS-CoV-2 con potencial de generar inmunidad cruzada con otros coronavirus humanos.

## **OBJETIVOS**

### **General**

Identificar los epítomos que pueden ser candidatos potenciales para generar inmunidad cruzada contra SARS-CoV-2 y otros coronavirus humanos de las proteínas estructurales S, M, N y E, así como de las accesorias ORF1ab, ORF3a, ORF6, ORF7a, ORF7b, ORF8 y ORF10.

### **Particulares**

1. Identificar los epítomos de las proteínas S, M, N, E, ORF1ab, ORF3a, ORF6, ORF7a, ORF7b, ORF8 y ORF10 de SARS-CoV-2 de células B y T que hayan sido reportados en la literatura.
2. Identificar, por medio de herramientas bioinformáticas, antígenos lineales de células B usando la secuencia de las proteínas S, M, N, E, ORF1ab, ORF3a, ORF6, ORF7a, ORF7b, ORF8 y ORF10 de SARS-CoV-2.
3. Identificar, por medio de herramientas bioinformáticas, epítomos de células T restringidos a MHC-I, MHC-II y epítomos promiscuos usando la secuencia de las proteínas S, M, N, E, ORF1ab, ORF3a, ORF6, ORF7a, ORF7b, ORF8 y ORF10 de SARS-CoV-2.

4. Identificar epítomos candidatos de las proteínas S, M, N, E, ORF1ab, ORF3a, ORF6, ORF7, ORF7b, ORF8 y ORF10 para generar inmunidad cruzada contra SARS-CoV-2 y otros coronavirus humanos.
5. Realizar modelos estructurales de los epítomos resultantes para cada una de las proteínas del SARS-CoV-2 analizadas en este estudio.

## MÉTODO

### Estrategia general

Como primer paso, de la base de datos se obtuvieron las secuencias de todas las proteínas de SARS-CoV-2 para la predicción de los epítomos. El segundo paso comprendió el análisis de dichas secuencias por medio de herramientas bioinformáticas para la predicción de epítomos de células B y T. Se siguió con la evaluación de los epítomos y se eligieron los epítomos potenciales a generar inmunidad cruzada contra la infección por SARS-CoV-2 y otros coronavirus humanos (Figura 8). A continuación, se describen a detalle cada uno de los pasos seguidos en la estrategia experimental.

### Recuperación de datos

La secuencia del genoma y las secuencias de las proteínas de SARS-CoV-2 (S, M, N, E, ORF1ab, ORF3a, ORF6, ORF7a, ORF7b, ORF8 y ORF10) se obtuvieron de la base de datos del Centro Nacional de Información Biotecnológica (NCBI, por sus siglas en inglés). Asimismo, se accedió a las secuencias del genoma de los coronavirus humanos HCoV-HKU1 (HKU), HCoV-NL63 (NL63), HCoV-OC43 (OC43), HCoV-229E (229E), MERS-CoV y SARS-CoV para cada una de las proteínas antes mencionadas a través de NCBI y UniProt (Recurso Proteico Universal, por sus siglas en inglés). En el caso de las proteínas ORF3a, ORF6, ORF7a, ORF7b y ORF10 se accedió únicamente a las secuencias del genoma de SARS-CoV, sólo las secuencias de ORF3a y ORF7a se obtuvieron de UniProt, el resto de NCBI; mientras que para la proteína ORF8 se accedió a las secuencias del genoma de SARS-CoV y MERS-CoV obtenidas de UniProt. Todos los números de acceso de las proteínas analizadas en este estudio se muestran en la Tabla 1:



**Tabla 1. Número de acceso de las proteínas de SARS-CoV-2 y de los coronavirus humanos.**

Proteína	Número de acceso						
	SARS-CoV-2	SARS	MERS	HKU	NL63	OC43	229E
S	QHR63290.2	AAU04646.1	AKN11075.1	YP_173238.1	YP_003767.1	YP_009555241.1	ABB90529.1
M	YP_009724393.1	AAU04638.1	YP_009047210.1	AYN64564.1	YP_003770.1	AAA45462.1	NP_073555.1
N	QIC53221.1	ABI96968.1	QFQ59594.1	AMM42397.1	AFV53152.1	QDH43730.1	NP_073556.1
E	YP_009724392.1	YP_009825054.1	YP_009047209.1	sp Q0ZME5.1	QED88042.1	AMK59679.1	sp P19741.1
ORF1ab	YP_009724389.1	NP_828849.7	tr A0A140AYW3	tr Q0ZJN1	tr A0A5B9BKB0	tr A0A1C8Z4S6	sp P0C6X1
ORF3a	YP_009724391.1	tr J9TEM7	-	-	-	-	-
ORF6	YP_009724394.1	sp P59634	-	-	-	-	-
ORF7a	YP_009724395.1	sp P59635	-	-	-	-	-
ORF7b	YP_009725318.1	tr Q7T6S2	-	-	-	-	-
ORF8	YP_009724396.1	sp Q80H93	tr A0A2D0Y3F8	-	-	-	-
ORF10	YP_009725255.1	tr Q6S8D9	-	-	-	-	-

\* Obtenidos de NCBI y UniProt

### Predicción de epítomos lineales de células B

En una primera etapa, se realizaron las predicciones de los epítomos de células B lineales de las proteínas S, M, N, E, ORF3a, ORF6, ORF7a, ORF7b, ORF8 y ORF10 de SARS-CoV-2 a través del sitio web IEDB, utilizando los algoritmos de predicción de epítomos lineales BepiPred V1.0 y V2.0. Cada uno de los epítomos se eligieron de acuerdo con los valores de unión de IC<sub>50</sub> de <50 (alta afinidad) y <500 (afinidad media).

### Predicción de epítomos de células T

Para la predicción de los epítomos de células T con restricción de MHC-I se usó el método NetMHCpan4.0 del recurso IEDB, esta predicción se realizó para 6 alelos HLA-A (HLA-A\*01:01, HLA-A\*02:01, HLA-A\*03:01, HLA-A\*11:01, HLA-A\*23:01, HLA-A\*24:02) y 6 alelos HLA-B (HLA-B\*07:02, HLA-B\*08:01, HLA-B\*35:01, HLA-B\*40:01, HLA-B\*44:02, HLA-B\*44:03). La selección de los epítomos con restricción MHC-I se basó en puntuaciones de afinidad IC<sub>50</sub> (<500) y puntuaciones de inmunogenicidad (>-1 y <1), valores que corresponden a una especificidad del 85%

y una sensibilidad del 30% y que se recomiendan en la estrategia de selección de péptidos potenciales.

Los péptidos con restricción MHC-II se predijeron mediante el uso del algoritmo 2.18 recomendado por IEDB. La predicción se realizó para 3 alelos HLA-DR (HLA-DRB1\*01:01, HLA-DRB1\*04:01, HLA-DRB1\*07:01), 8 alelos HLA-DP (HLA-DPA1\*01:03/DPB1\*03:01, HLA-DPA1\*02:01/DPB1\*02:01, HLA-DPA1\*02:02/DPB1\*02:02, HLA-DPA1\*03:01/DPB1\*23:01) y 6 alelos HLA-DQ (HLA-DQA1\*05:01/DQB1\*03:01, HLA-DQA1\*01:01/DQB1\*05:01, HLA-DQA1\*03:01/DQB1\*03:02) con un límite de longitud del péptido de 12 aa que es la longitud promedio que permite al núcleo del receptor del MHC-II para ser reconocidos por las células T y mantenerse unido a él y una mediana del percentil consenso del umbral de predicción  $\leq 20$  para lograr una alta sensibilidad para la predicción de epítomos. La selección del epítomo de células T se realizó con base a las puntuaciones de afinidad  $IC_{50}$  ( $< 500$ ).

La selección de HLA se realizó siguiendo las metodologías informadas previamente por Greenbaum<sup>58</sup> y Weiskopf<sup>59</sup>, y se utilizó para la predicción con el recurso TepiTool. Además, se utilizaron 4 alelos adicionales (1 para HLA-I y 3 para HLA-II) para la predicción, dado que éstos están presentes con una frecuencia alta en la población mexicana y también tienen una frecuencia significativa a nivel mundial (Tabla 2). Estos alelos se obtuvieron de la base de datos de frecuencias alélicas (<http://www.allelefrequencias.net/default.asp>), donde el criterio de búsqueda fue la región de América del Norte y todos los grupos de la población mexicana dentro de las capitales de todo el país.

**Tabla 2. Frecuencia en la población mundial de los alelos clase I y II utilizados en este estudio.**

Alelos	Frecuencia en la población mundial
<b>HLA Clase I</b>	
HLA-A*01:01	16.2
HLA-A*02:01	25.2
HLA-A*03:01	15.4
HLA-A*11:01	12.9
HLA-A*23:01	6.4
HLA-A*24:02	16.8
HLA-B*07:02	13.3
HLA-B*08:01	11.5
HLA-B*35:01	6.5
HLA-B*40:01	10.3
HLA-B*44:02	9.2
HLA-B*44:03	7.6
<b>HLA Clase II</b>	
HLA-DRB1*01:01	5.4
HLA-DRB1*04:01	6.2
HLA-DRB1*07:01	13.5
HLA-DQA1*05:01/HLA-DQB1*03:01	11.29
HLA-DQA1*05:01/HLA-DQB1*03:01	35.14
HLA-DQA1*03:01/HLA-DQB1*03:02	19.05
HLA-DQA1*04:01/HLA-DQB1*04:02	12.78
HLA-DQA1*01:01/HLA-DQB1*05:01	14.65
HLA-DQA1*01:02/HLA-DQB1*06:02	14.59
HLA-DPA1*02:01/HLA-DPB1*01:01	16.01
HLA-DPA1*01:03/HLA-DPB1*02:01	17.47
HLA-DPA1*01/HLA-DPB1*04:01	36.20
HLA-DPA1*03:01/HLA-DPB1*04:02	41.63
HLA-DPA1*02:01/HLA-DPB1*05:01	21.68

\* Los alelos con una frecuencia alta presentes en la población mexicana se muestran en rojo. Información obtenida de Allele frequency net database (AFND), 2020; Greenbaum J., *et al.*, 2011; Grifoni, *et al.*, 2020; Paul, *et al.*, 2016.

## Alineamiento múltiple

Se realizaron alineaciones de secuencias múltiples de las proteínas utilizando Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>), para luego utilizar el software Jalview versión 2.11.0. con el fin de visualizar de manera óptima cada una de estas secuencias. Posteriormente, para la identificación del porcentaje de identidad entre las secuencias de SARS-CoV-2 y los seis coronavirus humanos se usó la herramienta EMBOSS Needle ([https://www.ebi.ac.uk/Tools/psa/emboss\\_needle/](https://www.ebi.ac.uk/Tools/psa/emboss_needle/)), con la finalidad de elegir los epítomos prometedores que superaran el 50% de identidad entre ellos.

## Modelado estructural

Para generar una representación gráfica de los epítopos, se utilizó el modelado estructural de la glicoproteína S de SARS-CoV-2 de longitud completa (ID: 6VSB\_1\_1\_1) obtenido de la colección de sistemas proteicos de SARS-CoV-2 de la plataforma web CHARMM-GUI (<http://www.charmm-gui.org/docs/archive/covid19>) usando como guía lo presentado por Woo y colaboradores<sup>63</sup>. Por otro lado, las estructuras de la proteína M (ID: QHD43419), N (ID: QHD43423), E (ID: QHD43418), ORF3a (ID: QHD43417), ORF6 (ID: QHD43420), ORF7a (ID: QHD43421), ORF8 (ID: QHD43422) y ORF10 (ID: QHI42199) se obtuvieron a través de D-I-TASSER/C-I-TASSER (<https://zhanggroup.org/>). Las estructuras se construyeron en 3D y se analizaron utilizando el software PyMOL (Schrödinger LLC. Molecular Graphics System (pymol) versión 2.0.7 LLC, Nueva York, NY 2018). Finalmente, se utilizó la herramienta de búsqueda de alineación local básica en línea (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) para evaluar la posición de los péptidos predichos y reportados en la secuencia de las proteínas mostradas y de esta manera poder seleccionar cada uno de estos.

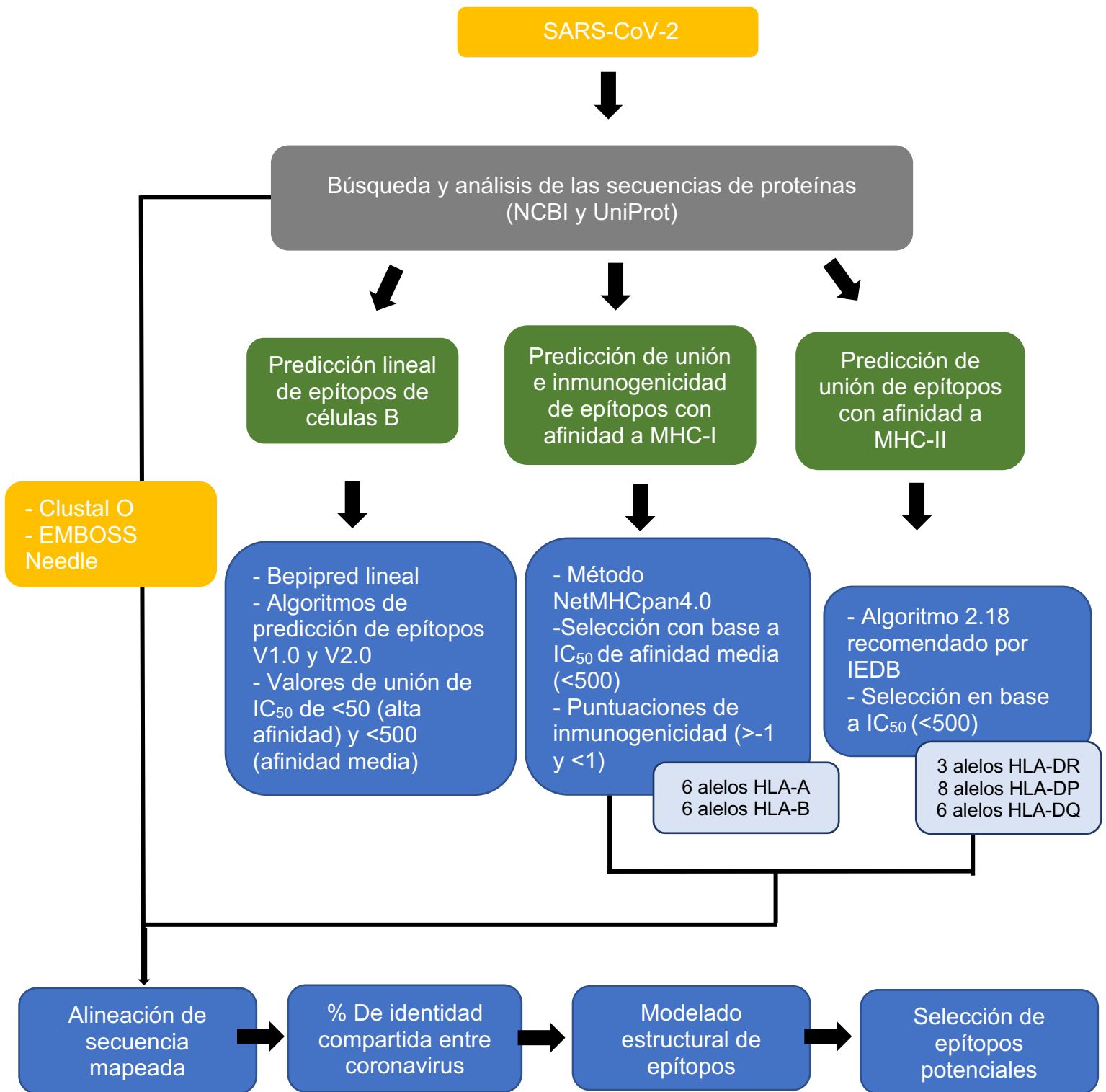


Figura 8. Flujo de trabajo del estudio.

## RESULTADOS

Utilizando herramientas basadas en la web, se llevó a cabo un análisis bioinformático de las secuencias de las proteínas de SARS-CoV-2, con el fin de identificar nuevos epítomos de células B y T con potencial para ser utilizados como dianas de vacunas. A continuación, se describen los distintos resultados obtenidos:

### **Predicción de epítomos lineales de células B**

Mediante el uso de BepiPred 1.0 y 2.0, se predijeron 459 epítomos en función de su puntuación y accesibilidad a la superficie. En total, se compilaron 235 epítomos potenciales de células B (de este estudio e informes anteriores) de los cuales 64 fueron predichos en este estudio y reportados por otros grupos y 171 fueron nuevos epítomos predichos en este estudio. Dentro de los epítomos predichos en este estudio, 17 están localizados en la proteína S, 1 en M, 3 en N, 133 en la proteína ORF1ab, 8 en ORF3a, 1 en ORF6, 3 en ORF7a, 1 en ORF7b y 4 en ORF8. El único epítomo predicho para la proteína E, ya estaba reportado en otros estudios, mientras que para la proteína ORF10 no se obtuvo ningún resultado de acuerdo con la puntuación y accesibilidad a la superficie.

Todos los epítomos reportados en otros estudios para cada una de las proteínas se muestran en la Tabla 3, en tanto que los epítomos predichos en este estudio se muestran en el Anexo 1, Tabla 7. Aquellos epítomos predichos en este estudio que coincidían con los epítomos reportados o tenían residuos superpuestos se muestran en negritas en la Tabla 3. Asimismo, en cada celda de la tabla, se muestran de manera puntual las distintas secuencias de los epítomos que los autores han reportado en sus respectivas investigaciones.

**Tabla 3. Epítomos reportados en la literatura de células B lineales.**

Proteína	Secuencia	Posición	Fuente
S	1) MFVFLVLLPLVSSQC	1-15	(64), (65), (66)
	2) VLLPLVSSQCVNLTTRTQLPPAYTN		
	3) CVNLTTRTQLPPAYTN		
	RSSVLHSTQD	44-53	(65)
	1) NVTWFHAIHVSGTNGT	61-73	(64), (65), En este estudio
	2) VTWFHAIHV <b>SGTNGTKRFDN</b>		
	VYFASTEKSNII	90-101	(65)
	GTTLDSKTQSLIVNNATNVVIKVC	107-131	(65)
	YYHKNNKSWMESEFRVYSSANNCTFEYVVSQPFLM	144-177	(65)
	<b>IYSKHTPIN</b>	203-211	(65), En este estudio
	1) TRFQTLALHRSYLTSGDSSSGW	236-286	(65), (66)
	2) RSYLTSGDSSSGWTAGAAAY		
	3) LTPGDSSSGWTAGAAAYVGYLQPRTFLLKYN		
	4) DSSSGWTAGAAAYV		
	5) SSSSGWTAGAAAYVG		
	1) DAVDCALDPLSETKCTLKSFTVEKGIYQTSN	287-317	(65), (66), (67)
	2) DAVDCSQNPLAELKCSVKSFEIDKGIYQTSNF		
	3) KGIYQTSN		
	4) YQTSNFRVQP		
	SFSTFKCYGVSPTKLNDL	373-390	(64)
	<b>NNLDSKVGGNINY</b>	439-451	(66), En este estudio
	LFRKSNLKPFERDISTEIQAGST	455-478	(66), (75)
	1) ELLHAPAT <b>VC</b> PKKSTNLVK	516-535	(66), (67), En este estudio
	2) VCGPKKSTNLVKNKCVNFNGLTGTGVL <b>TESNKKFLPFQQGRDIADTTDA</b>		
	3) <b>VRDPQ</b> TLEILDITPCSFGGVSVI		
	GAGICASY	667-674	(67)
<b>AYTMSLGAENSVAYSN</b>	694-709	(66), En este estudio	
GSFCTQLN	757-764	(67)	
<b>FIEDLLFNKVTLADAGF</b>	817-833	(67), En este estudio	
GAALQIPFAMQMAYRFN	891-907	(67)	
QALNTLVKQLSSNFGAI	957-973	(67)	
1) EAEVQIDRLITGRLQSL	988-1004	(67)	
2) RLITGRLQSLQTYVTQQ			
<b>GQSKRVDFC</b>	1035-1043	(66), En este estudio	

	1) VNNTVYDPLQPELDSFKEELDKYF <b>KNHTSPDVLGD</b> ISGIN 2) TVYDPLQPELDSFKEEL	1132-1173	(66), (67), En este estudio
	1) CKFDEDDSEPVKGVKLYHT 2) SCCKFDEDDSEPVKGVKLYHT	1254-1273	(66), (67)
M	1) MADSNGTITVEELKKLLEQWNLVI 2) LEQWNLVIGFLFLT PVTLACFVLAAYR GGIAIAMAQLVGLM <b>LLESELVIGAVILRGHLRI</b>	1-24 59-72 78-91 133-151	(61), (66), (68) (68) (68) (61), En este estudio
	1) GRCDIKDLPKEITVATSR 2) ITVATSRT 3) PKEITVATSRTLSYYKL 4) TSRTLSYYKLGASQRV 5) KLGASQRVAGDS	157-174	(66), (67)
N	1) NGPQNQRNAPRIT <b>FGGPSDSTG</b> SNQNGERSGARSKQRRPQGLPNN 2) RPQGLPNNTASWFTALTQHGK 3) RRPQGLPNNTASWFT 4) HGKEDLKFPRGQGVPIINTNSSPDDQIGYYRRATR <b>RIRGGDGKMKDLS</b>	4-48 137-157	(61), (66), (67), En este estudio (61), (66), (67)
	1) GALNTPKDHIGTRNPANNAI 2) NNAAITVLQLPQGTTLPKGF 3) QLPQGTTLPKGFYAEGR 4) QLPQGTTLPKGFYAEGRGGSQ 5) TLPKGFYAEGRGGSQASSRSSSRNSSRNSTPGSSRGTSARMAGNGGD 6) VLQLPQGTTLPKGFYAEGRGGSQASSRSSSRNSSRNSTPGSSRGTSARMAGNGGD 7) MAGNGGD <b>1) LNQLESKMSGKGGQQQQQT</b> VTKKSAEASKKPRQKRTATK 2) RRGPEQTQGNFGDQELIRQGTDYK 3) KHWPQIAQFAPSASAFF 4) FFGMSRIGMEVTPSGTW 5) YNVTQAFGRRGPEQTQGNF	227-266	(66), (67), En este estudio
E	YVYSRVKLNSSRVPD RVKN	57-72 61-64	(66) (67)

\* En negrita se muestran los epítomos predichos en este estudio que coincidían con los reportados

Adicionalmente, se encontraron en la literatura 26 epítomos de células B lineales, para los cuales, su capacidad de ser reconocidos para montar una respuesta inmunitaria ha quedado comprobada de forma experimental (Anexo 1, Tabla 9). De estos epítomos, 19 son de la proteína S, 2 de la proteína M y 5 de la proteína N. Cabe destacar que 7 de estos, todos de la proteína S, tienen residuos



superpuestos que coinciden con epítomos predichos en este estudio, los cuales se muestran en negritas en la Tabla 9.

Por otro lado, la alineación de secuencia múltiple de las proteínas del SARS-CoV-2 contra las secuencias de las proteínas de otros coronavirus humanos (SARS-CoV, MERS-CoV, HKU1, NL63, OC43 y 229E) permitió identificar 181 epítomos conservados entre estos coronavirus que podrían ser objetivo de anticuerpos de reactividad cruzada. Cabe señalar que, para las secuencias de las proteínas ORF3a, ORF6, ORF7a y ORF7b, sólo fue posible compararlas con la del SARS-CoV, mientras que la de la proteína ORF8 se comparó con la secuencia del SARS-CoV y MERS-CoV, debido a que al momento de realizar este estudio no se encontraron disponibles las secuencias de los coronavirus faltantes en las bases de datos GenBank y UniProt.

Siguiendo esta estrategia, se identificaron 157 epítomos diferentes que comparten cierta identidad entre los siete coronavirus (Anexo 2, Tabla 10). Al menos, 24 péptidos mostraron un porcentaje de identidad superior al 50% entre el SARS-CoV-2 y todas las demás secuencias de coronavirus, la mayoría obtenidos de la proteína ORF1ab (Tabla 4). Como era de esperar, los epítomos entre el SARS-CoV y el SARS-CoV-2 compartieron los porcentajes de identidad más altos entre los coronavirus.

**Tabla 4. Epítomos de células B conservados, con identidad superior al 50% entre el SARS-CoV-2 y seis coronavirus humanos.**

Proteína	SARS-CoV-2	SARS	MERS	229E	HKU1	NL63	OC43
S	GQSKRVDFC	100	66.7	66.7	55.6	55.6	55.6
ORF1ab	LTPCGTGTSTD	100	54.5	54.5	63.6	54.5	63.6
	FPPTSFG	100	57.1	57.1	57.1	57.1	57.1
	VKPGNFNK	100	75	87.5	75	75	87.5
	GNAAIS	100	83.3	66.7	83.3	50	83.3
	MSYEDQ	100	66.7	66.7	50	66.7	50
	SDVENP	83.3	66.7	50	66.7	66.7	66.7
	YPKCDRA	100	100	100	100	100	100
	YVKPGGTSSGDATTAY	100	100	87.5	88.2	81.2	88.2
	ETDLTKGPHEF	100	90.9	72.7	72.7	90.9	63.6
	DDYVYLPYPDPSRI	100	78.6	71.4	92.9	78.6	78.6
	PLTKHPNQEY	100	70	70	60	80	60
	CNAPGCDVT	100	77.8	55.6	77.8	66.7	88.9
	VGKPRPLN	100	66.7	66.7	66.7	66.7	66.7
	STLQGPPGTGKS	100	83.3	75	83.3	75	83.3
	GDPAQLPAP	100	100	88.9	100	88.9	100
	SSAINR	100	100	66.7	66.7	66.7	66.7
	YNSQNA	100	100	83.3	83.3	83.3	83.3
	TQTVDSQSEY	100	100	91.7	91.7	91.7	91.7
	HPTQAPTHLS	100	50	50	50	60	50
	IGNPKA	100	83.3	83.3	83.3	83.3	83.3
LPGCDGGS	100	100	62.5	87.5	62.5	87.5	
AGSDKGVAPGT	100	90.9	90.9	81.8	100	81.8	
DFVSDADST		90.9	100	66.7	55.6	55.6	77.8

\* La tabla está coloreada según el porcentaje de conservación (más oscuro cuanto mayor es el porcentaje de identidad)

Para visualizar la posición de los epítomos compilados se utilizó un modelo 3D optimizado de las proteínas S, M, N, E, ORF3a, ORF6, ORF7a, ORF8 y ORF10 del SARS-CoV-2 (Figura 9), se representaron los epítomos reportados en la literatura (en color magenta), los epítomos que se obtuvieron de las predicciones en este estudio (en color rosa) y finalmente los epítomos que ya fueron evidenciados de ser capaces de ser reconocidos para montar una respuesta inmunitaria de manera experimental (en color morado). Los epítomos de células B descritos en la literatura se distribuyen en todos los dominios de la glicoproteína de espícula de manera similar a los epítomos predichos en este estudio. El modelo trimérico muestra que los epítomos de células B lineales que ya fueron evidenciados de forma experimental, no son tan abundantes en comparación con los reportados en estudios bioinformáticos realizados por otros grupos y aquellos que fueron predichos en este estudio. La mayoría de estos epítomos de la glicoproteína S se

localizan en las regiones que comprenden la subunidad S1 que contiene el dominio NTD, el dominio RBD y el RBM (Figura 9-A). Para la proteína M, los epítomos reportados y predichos en este estudio se pueden observar en la zona del dominio C-terminal y en el dominio N-terminal (Figura 9-B). Los epítomos de la proteína de la nucleocápside se localizaron a lo largo de la proteína desde el dominio N-terminal hasta el dominio C-terminal, sin embargo, la concentración de la mayoría de los epítomos se pudo encontrar en el dominio de unión al RNA y en el dominio de dimerización (Figura 9-C). En el caso de la proteína E todos los epítomos reportados se situaron en el dominio C-terminal (Figura 9-D).

Es de destacar que, a pesar de no contar con todos los modelos estructurales cristalizados de los ORF's de SARS-CoV-2, todos los epítomos de células B lineales predichos en este trabajo, no han sido reportados en otros estudios. De forma específica, los epítomos de la proteína ORF3a se encontraron en su mayoría en el dominio citosólico de las láminas  $\beta$  plegadas (Figura 9-E), mientras que de la proteína ORF6 los epítomos se observaron en una de las estructuras  $\alpha$  hélices (Figura 9-F). Para la proteína ORF7a los epítomos se localizaron a lo largo de toda la proteína, en su mayoría en el dominio N-terminal (Figura 9-G), al igual que la proteína ORF8 cuyos epítomos se situaron también en la región del dominio N-terminal (Figura 9-H).

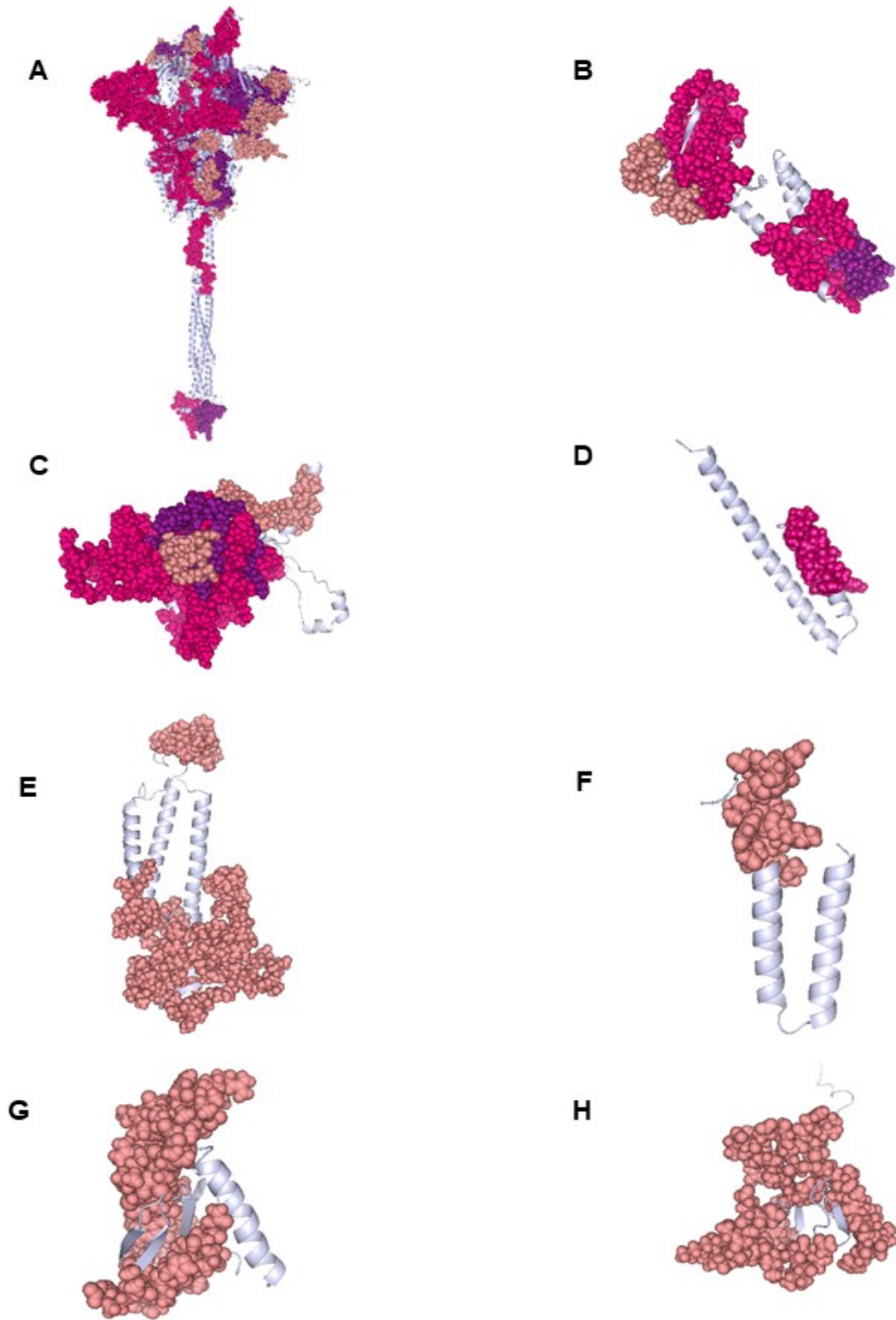


Figura 9. Epítomos lineales de células B de las proteínas S (A), M (B), N (C), E (D), ORF3a (E), ORF6 (F), ORF7a (G) y ORF8 (H). En magenta se representan los epítomos reportados en la

literatura, en rosa los epítomos que se obtuvieron de las predicciones en este estudio y en morado los epítomos analizados experimentalmente y capaces de ser reconocidos para montar una respuesta inmunitaria.

### **Predicción de epítomos de células T**

Mediante el método recomendado por IEDB, se obtuvieron un total de 245 posibles epítomos de células T, de los cuales 120 han predicho la unión a MHC-I, 87 han predicho la unión a MHC-II y 38 fueron epítomos promiscuos. Asimismo, se desglosaron todas las secuencias que han sido descritas por otros autores y en negrita se señalaron los epítomos predichos en este estudio que coincidían con los reportados en la literatura. Por último, para los epítomos promiscuos se señaló en negrita el *core* del péptido (Tabla 5).

En tanto que los epítomos predichos en este estudio se obtuvieron un total de 391 posibles epítomos de células T, de los cuales 327 han predicho la unión a MHC-I, 30 han predicho la unión a MHC-II y 34 se pronostica que sean epítomos promiscuos de células T (Anexo 1, Tabla 8).

Cabe señalar que no se predijeron epítomos en este estudio de las proteínas E, ORF7a, ORF7b, ORF8 y ORF10 con restricción a MHC-I y de las proteínas ORF7a y ORF7b con restricción a MHC-II, así como de los epítomos promiscuos de la proteína E de los cuales tampoco se obtuvo ningún resultado. La mayoría de estos epítomos se predijeron para la proteína ORF1ab, con un total de 300 epítomos, impulsado por el gran tamaño de ésta dentro del proteoma del SARS-CoV-2.

Como en el caso de los epítomos lineales de las células B, aquellos epítomos predichos en este estudio que coinciden con los epítomos reportados o tienen residuos superpuestos, se muestran en negritas en la Tabla 5. De igual forma, para los epítomos promiscuos, se señalan en negrita los aminoácidos que se mantienen sin cambios y se comparten en cada péptido. Por último, en cada celda de la tabla se muestran de manera puntual las distintas secuencias de los epítomos que los autores han reportado en sus investigaciones.

**Tabla 5. Epítomos reportados en la literatura de células T restringidos a MHC-I, MHC-II y promiscuos.**

Proteína	Secuencia	Posición	MHC	HLA	Fuente
S	<b>GVYFASTEKSNI</b>	89-100	I	HLA-A*03:01, HLA-A*11:01	(64), En este estudio
	WTAGAAAYVGYLQPRTFLK	256-278	I	HLA-A*02:01, HLA-A*02:06, HLA-B*08:01, HLA-B*15:01	(64)
	CVADYSVLY	361-369	I	HLA-A*01:01, HLA-A*26:01, HLA-A*29:02, HLA-A*30:02	(67)
	KLPDDFTGCV	424-433	I	HLA-A*02:01	(61)
	NYNLYRLF	448-456	I	HLA-A*23:01, HLA-A*24:02	(61), (66)
	<b>GYQPYRVVLSF</b>	504-515	I	HLA-A*23:01, HLA-A*24:02, HLA-A*29:02, HLA-B*08:01	(67), En este estudio
	IGAGICASYQTQ	666-677	I	HLA-A*29:02, HLA-A*30:02	(67)
	LLLQYGSFCTQLNRALT	752-768	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*68:01	(67)
	SFIEDLLFNK	816-825	I	HLA-A*11:01, HLA-A*31:01, HLA-A*33:01, HLA-A*68:01	(67)
	TQNVLYENQK	912-921	I	HLA-A*11:01, HLA-A*68:01	(67)
	RVDFCGKGY	1039-1047	I	HLA-A*01:01, HLA-A*03:01, HLA-A*30:02, HLA-B*15:01, HLA-B*27:05	(67)
	<b>APHGVVFLHV</b>	1056-1065	I	HLA-B*07:02, HLA-B*35:01, HLA-B*53:01, HLA-B*54:01	(67), En este estudio
	IITDNTFV	1114-1122	I	HLA-A*02:01	(69)
	VYDPLQPELDSFKEELDKY	1137-1155	I	HLA-A*01:01, HLA-A*24:02, HLA-A*26:01, HLA-A*29:02, HLA-A*30:02	(67)
	<b>RLNEVAKNL</b>	1185-1193	I	HLA-A*02:01	(61), (67), (70), En este estudio
	NESLIDL	1194-1200	I	HLA-A*02:01	(61), (67)
	CMTSCCCLK	1236-1245	I	HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*33:01, HLA-A*68:01	(67)
DEDDSEPVLGKVKL	1257-1270	I	HLA-B*40:01, HLA-B*40:02, HLA-B * 44:02, HLA-B*44:03	(67)	
M	1) KLEQWNLV 2) LEQWNLVIGF	15-23	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*03:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-A*26:01, HLA-A*29:02, HLA-A*32:01, HLA-A*68:01, HLA-A*68:02, HLA-B*15:01, HLA-B*18:01, HLA-B*40:01, HLA-B*40:02, HLA-B*44:02, HLA-B*44:03, HLA-B*45:01, HLA-B*57:01, HLA-B*58:01	(66), (67), (69), (71)
	1) YFIASFRLF 2) RTRSMWSF 3) MWSFNPETNI	95-103	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*23:01, HLA-A*24:02, HLA-A*30:01, HLA-A*30:02, HLA-A*32:01, HLA-B*14:02, HLA-B*57:01, HLA-B*58:01, HLA-C*06:02, HLA-C*07:01	(66), (67), (71)

	1) SELVIGAVI 2) LVIGAVILR	136-144	I	HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*33:01, HLA-A*68:01, HLA-B*40:01, HLA-B*40:02, HLA-B*44:03, HLA-B*45:01, HLA-B*18:01, HLA-B*44:02	(67), (71)
	1) DSGFAAYSRY 2) RIGNYKL	190-199	I	HLA-A*01:01, HLA-A*23:01, HLA-A*24:02, HLA-A*26:01, HLA-A*30:02	(67), (75)
N	FPRGQGVPI	66-74	I	HLA-A*02:01, HLA-B*07:02, HLA-B*08:01, HLA-B*35:01, HLA-B*51:01, HLA-B*53:01, HLA-B*54:01	(66), (67)
	LSPRWYFYY	104-112	I	HLA-A*01:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-A*29:02, HLA-A*30:02, HLA-A*31:01, HLA-B*07:02, HLA-B*51:01, HLA-B*53:01, HLA-B*54:01	(67)
	1) ATEGALNTPK 2) ALNTPKDHI	134-143	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*68:01	(66), (67)
	1) AEGSRGGSQA 2) SQASSRSSSR	173-182	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*26:01, HLA-A*29:02, HLA-A*30:02, HLA-A*31:01, HLA-A*33:01, HLA-A*68:01, HLA-B*45:01	(67)
	1) GDAALALLL 2) LLLLDRLNQL 3) RLNQLESKM	215-224	I	HLA-A*02:01, HLA-B*40:01	(61), (67), (69)
	QQQQGQTVTK	239-248	I	HLA-A*11:01	(67)
	1) NFKDQVILL 2) ILLNKHIDA	345-353	I	HLA-A*02:01, HLA-A*24:02	(61), (67)
E	SEETGTLIV	6-14	I	HLA-B*45:01, HLA-B*40:01	(67)
	RLCAYCCNI	38-46	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67)
ORF1ab	<b>WSMATYYLF</b>	900-908	I	HLA-A*23:01, HLA-A*24:02, HLA-B*58:01	(66), En este estudio
	LLSAGIFGA	1148-1156	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67)
	<b>ILGTVSWNL</b>	1367-1375	I	HLA-A*02:01, HLA-B*58:01	(67), En este estudio
	ETISLAGSY	1503-1511	I	HLA-A*26:01	(66)
	<b>FLGRYMSAL</b>	1642-1650	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06	(67), En este estudio
	<b>GEEANFCAL</b>	1705-1713	I	HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	(66), En este estudio
	VMYMGTLISY	1768-1776	I	HLA-B*15:01	(66)
	<b>KPNELSRVLGLK</b>	2109-2120	I	HLA-A*03:01, HLA-B*07:02	(66), En este estudio
	LAAVNSVPW	2127-2135	I	HLA-B*58:01	(66)
	QLMCQPILLL	2563-2572	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67)
	TPRDLGACI	2684-2692	I	HLA-B*07:02, HLA-B*51:01, HLA-B*54:01	(67)
	<b>KSHNIALIWNVK</b>	2705-2716	I	HLA-A*11:01, HLA-B*58:01	(66), En este estudio

HFLPRVFAV	2883-2892	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-B*07:02, HLA-B*08:01	(67), En este estudio
1) FLAHIQWMV 2) VSFLAHIQW	3122-3130	I	HLA-A*02:01, HLA-B*58:01, HLA-B*08:01, HLA-B*35:01	(66), En este estudio
VMFTPLVPF	3130-3138	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*15:01, HLA-B*35:01	(66), En este estudio
FLNGSCGSV	3403-3411	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67)
TVNVLAWLYAAV	3464-3475	I	HLA-A*01:01, HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*11:01, HLA-A*68:02, HLA-B*35:01	(67), En este estudio
FLNRFITTL	3482-3490	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67), En este estudio
RTIKGTHHW	3574-3582	I	HLA-B*58:01	(66)
TLMNVLTLYV	3710-3719	I	HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-B*35:01	(66), En este estudio
FLARGIVFM	3753-3761	I	HLA-A*02:01	(66), En este estudio
1) TLGVYDYL 2) GVDYDLVST	3807-3815	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67), En este estudio
FRYMNSQGL	3820-3828	I	HLA-C*06:02	(67)
KLWAQCVQL	3886-3894	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*03:01, HLA-A*31:01	(67)
SLLSVLLSM	3913-3921	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06	(67)
AMQTMFLTMRK	4028-4039	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67), En este estudio
1) FTYASALWEI 2) ALWEIQVV	4089-4098	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*03:01, HLA-A*11:01, HLA-A*68:02, HLA-A*69:01, HLA-B*15:01	(67)
TVLSFCAFAV	4265-4274	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67)
1) YTMADLVYA 2) TMADLVYAL	4514-4522	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(66), (67)
ILHCANFNV	4699-4707	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*31:01, HLA-A*68:02, HLA-A*69:01	(67)
KIFVDGVPFVV	4724-4734	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-B*35:01	(67), En este estudio
MHAASGNLL	4772-4780	I	HLA-A*39:01	(66)
ISDYDYRY	4842-4850	I	HLA-A*01:01, HLA-A*29:02, HLA-A*30:02	(67)



<b>RQLLFVVEV</b>	4859-4867	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67), En este estudio
RLYYDSMSY	4905-4913	I	HLA-A*03:01, HLA-A*11:01, HLA-B*15:01	(66), (67)
<b>ATVVGTSK</b>	4977-4985	I	HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*68:01	(67), En este estudio
MVMCGGSLYV	5058-5067	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67)
1) <b>DTDFVNEFY</b> 2) <b>FVNEFYAYL</b>	5130-5138	I	HLA-A*01:01, HLA-A*02:01	(66), En este estudio
YLRKHFSMM	5140-5148	I	HLA-B*08:01	(66)
RILGAGCFV	5228-5236	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06	(67)
LVLSVNPYV	5365-5373	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67)
<b>KLFAAETLK</b>	5455-5463	I	HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*68:01	(67), En este estudio
<b>KLSYGIATVR</b>	5470-5479	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*11:01, HLA-A*68:02	(67), En este estudio
<b>VVYRGTTY</b>	5533-5541	I	HLA-A*03:01, HLA-A*11:01, HLA-B*15:01, HLA-B*35:01, HLA-B*58:01	(67), En este estudio
KLNVGDYFV	5542-5550	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67)
TLVPQEHYV	5563-5571	I	HLA-A*02:01, HLA-A*69:01	(67)
1) <b>SSNVANYQK</b> 2) <b>YQKVGMMQKY</b>	5587-5595	I	HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*68:01, HLA-B*15:01, HLA-B*27:05	(67)
<b>SHFAIGLALYY</b>	5613-5623	I	HLA-A*01:01, HLA-A*39:01, HLA-B*35:01	(66), En este estudio
<b>LEQYVFCTVNAL</b>	5676-5687	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-B*35:01, HLA-B*40:01	(67), En este estudio
<b>IVDTVSALVYDNK</b>	5772-5784	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*11:01, HLA-A*68:02	(67), En este estudio
<b>IPRRNVATL</b>	5916-5924	I	HLA-A*02:01, HLA-B*07:02, HLA-B*35:01, HLA-B*08:01	(66), (67), En este estudio
LQLGFSTGV	6032-6040	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67)
<b>LMYKGLPWNVVRI</b>	6077-6089	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*23:01, HLA-A*24:02	(67), En este estudio
<b>VFVLWAHGFEL</b>	6107-6117	I	HLA-A*01:01, HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*03:01, HLA-A*23:01, HLA-A*24:02, HLA-A*31:01, HLA-A*68:02, HLA-A*69:01, HLA-B*08:01, HLA-B*15:01, HLA-B*58:01	(67), En este estudio

	<b>ASCDAIMTRCLAV</b>	6195-6207	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*11:01, HLA-A*68:02	(67), En este estudio
	ALLADKFPV	6245-6253	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06	(67)
	KAYKIEELF	6284-6290	I	HLA-B*58:01	(66)
	<b>YLDAYNMMI</b>	6419-6427	I	HLA-A*02:01	(66), En este estudio
	<b>KQFDTYNLW</b>	6437-6445	I	HLA-A*23:01, HLA-A*24:02, HLA-A*31:01, HLA-B*15:01, HLA-B*58:01	(67), En este estudio
	NLWNTFTRL	6443-6451	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67)
	<b>TTLPVNVAFEL</b>	6499-6509	I	HLA-A*02:01, HLA-A*69:01, HLA-B*07:02, HLA-B*15:01, HLA-B*35:01, HLA-B*58:01	(67), En este estudio
	1) LLLDDFVEI 2) LLDDFVEII	6749-6757	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06	(66), (67)
	<b>EISFMLWCKDGHV</b>	6778-6790	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*11:01, HLA-A*68:02	(67), En este estudio
	GVAMPNLYK	6806-6814	I	HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*68:01	(67)
	KMQRMLEK	6814-6822	I	HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*68:01	(67)
	<b>YLNTLTLAV</b>	6851-6859	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(66), (67), En este estudio
	<b>RQWLPTGTLV</b>	6884-6894	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-B*40:01	(67), En este estudio
	TLIGDCATV	6908-6916	I	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02	(67)
ORF3a	<b>LLYDANYFLCW</b>	139-149	I	HLA-A*02:01, HLA-B*44:02, HLA-B*44:03	(66), En este estudio
ORF7a	1) YEGNSPFHPL 2) HPLADNKFAL	40-49	I	HLA-B*40:01, HLA-B*44:03, HLA-B*45:01	(67)
	KLFIRQEEV	85-93	I	HLA-A*02:01, HLA-A*31:01, HLA-B*08:01	(67)
S	<b>RGWIFGTTLDSKTQSL</b>	102-116	II	HLA-DRB1*04:01, HLA-DRB1*07:01	(61), En este estudio
	QPFLMDLEGKQGN	173-185	II	HLA-DRB1*04:01	(61)
	TRFQTLALHRSYLTPGDSSSG W	236-258	II	HLA-DRB1*04:01	(61)
	KSFTVEKGIYQTSNFRVQ	304-321	II	HLA-DRB1*04:01, HLA-DRB1*07:01	(61)
	<b>SASFSTFKCYGVSPTKL</b>	371-385	II	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DR8, HLA-DQA1*05:01, HLA-DQB1*03:01	(61), En este estudio
	YLYRLFRKSNLKPFERDISTEI	451-472	II	HLA-DRB1*04:01	(61)
	<b>CVNFNFNGLTGTGVL</b>	538-552	II	HLA-DRB1*01:01, HLA-DQB1*03:01, HLA-DRB1*07:01, HLA-DQA1*05:01	(67), En este estudio

	<b>PCSFGGVSVITPGTN</b>	589-603	II	HLA-DRB1*01:01, HLA-DQB1*03:01, HLA-DQA1*05:01	(67), En este estudio
	TECSNLLLQYGSFCTQL	747-761	II	HLA-DR8	(67)
	<b>VKQIYKTPPIKDFGGFNF</b>	785-802	II	HLA-DRB1*04:01, HLA-DPA1*02:02, HLA-DPB1*02:02	(61), En este estudio
	EDLLFNKVTLADAGF	819-833	II	HLA-DRB1*01:01	(67)
	AQKFNGTLVLPPLT	852-866	II	HLA-DRB1*01:01	(67)
	<b>DSLSTASALGKLQDVV</b>	936-952	II	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DQA1*05:01, HLA-DQB1*03:01	(66), En este estudio
	GVVFLHVTY	1059-1067	II	HLA-DPA1*02:02, HLA-DPB1*02:02	(67)
	HWFVTQRNFYEPQII	1101-1115	II	HLA-DRB1*04:01	(61)
	LGDISGINASVVNIQ	1166-1180	II	HLA-DRB1*01:01	(67)
	DDSEPVKGVKLHYT	1259-1273	II	HLA-DRB1*01:01	(67)
M	VIGFLFLTWICLLQFAYANRNR FLYIIKLIFLWLLWPVTLACF	23-65	II	HLA-DPA1*01:03, HLA-DPB1*02:01, HLA-DPB1*04:02, HLA-DPB1*05:01, HLA-DPB1*09:01, HLA-DQA1*01:01, HLA-DQA1*01:02, HLA-DQA1*01:03, HLA-DQA1*03:01, HLA-DQB1*03:02, HLA-DQB1*05:01, HLA-DQB1*06:01, HLA-DQB1*06:02	(67), (69), (71)
	1) LRGHLRIAGHHLGR 2) ILRGHLRIAGHHLGR	145-158	II	HLA-DRB1*01:01	(66)
	TLSYYKLGASQRVAG	175-189	II	HLA-DRB1*01:01	(66)
N	PDDQIGYYRRATRRIRGGDGK	80-100	II	HLA-DPA1*01:03, HLA-DPA1*02:01, HLA-DPB1*02:01, HLA-DPB1*04:02, HLA-DPB1*05:01, HLA-DPB1*09:01	(69)
	<b>GIIVVATEGALNTPK</b>	129-143	II	HLA-DRB1*01:01	(66), En este estudio
	<b>GTRNPANNAIIVLQLPQGTTL</b>	147-167	II	HLA-DRB1*01:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPB1*05:01, HLA-DPB1*09:01, HLA-DQA1*01:01, HLA-DQA1*01:02, HLA-DQA1*05:01, HLA-DQB1*03:01, HLA-DQB1*06:02	(67), (69), En este estudio
	<b>RMAGNGGDAALALLLDRLN QLESK</b>	209-233	II	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DQA1*01:03, HLA-DQA1*03:01, HLADQA1*05:01, HLA-DQB1*03:01,	(61), (69), En este estudio

				HLA-DQB1*03:02, HLA-DQB1*03:03, HLA-DQB1*04:01, HLA-DQB1*06:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:02, HLA-DPB1*23:01	
	GQTVTKKSAAEASKK	243-257	II	HLA-DRB1*01:01, HLA-DQA1*05:01, HLADQB1*03:01	(67)
	QKKQQTVTLLPAADLDDFS	386-404	II	HLA-DRB1*09:01, HLA-DRB1*13:02	(69)
E	1) LVKPSFYVYSRVKLN 2) VKPSFYVYSRVKLN 3) KPSFYVYSRVKLNLS 4) PSFYVYSRVKLNSS 5) SFYVYSRVKLNSSR 6) FYVYSRVKLNSSRV	51-65	II	HLA-DRB1*01:01, HLA-DQA1*01:01, HLA-DPA1*01:03, HLA-DPA1*02:02, HLA-DPB1*02:01, HLA-DPB1*04:02, HLA-DPB1*05:01, HLA-DPB1*06:01, HLA-DPB1*09:01	(69), En este estudio
ORF1ab	KPLEFGATSAALQPE	958-972	II	HLA-DQA1*02:01, HLA-DQB1*03:01	(66)
	ETKFLTENLLLYIDI	1245-1259	II	HLA-DPA1*03:01, HLA-DPB1*04:02	(66)
	KTTVASLINTLNDLN	1428-1442	II	HLA-DQA1*01:02, HLA-DQB1*05:01	(66)
	1) ESPFVMSAPPAQYE 2) PFVMSAPPAQYELK 3) SPFVMSAPPAQYEL	1801-1815	II	HLA-DRB1*01:01, HLA-DQA1*01:02, HLA-DQB1*05:01	(66)
	ASRELKVTFFPDNG	1951-1965	II	HLA-DPA1*01:03, HLA-DPB1*06:01	(66)
	1) AALGVLMNSNLGMPSY 2) ALGVLMNSNLGMPSYC 3) GSLIYSTAALGVLM	2248-2262	II	HLA-DQA1*01:02, HLA-DQA1*02:01, HLA-DQB1*03:01, HLA-DQB1*05:01	(66)
	1) SHFVNLNLRANNTK 2) HSLSHFVNLDNLRAN	2519-2533	II	HLA-DRB1*01:01, HLA-DRB1*04:04	(66)
	NVLSTFISAARQGFV	2628-2642	II	HLA-DQA1*02:01, HLA-DQB1*03:01	(66)
	EYTDFAATSACVLAEE	2904-2918	II	HLA-DQA1*02:01, HLA-DQB1*03:01	(66)
	AVNLLTNMFTPLIQP	3023-3037	II	HLA-DPA1*01:03, HLA-DPB1*04:01	(66)
	1) ISASIVAGGIVAIVV 2) DISASIVAGGIVAIV	3043-3057	II	HLA-DQA1*02:01, HLA-DQB1*03:01	(66)
	1) TPVYSFLPGVYSVIY 2) SVIYLYLTFYLTNDV 3) VYSVIYLYLTFYLTN	3095-3109	II	HLA-DRB1*01:01, HLA-DPA1*01:03, HLA-DPB1*06:01	(66)
	1) QWSLFFFLYENAF 2) SLFFFLYENAF 3) LFFFLYENAF	3599-3613	II	HLA-DPA1*01:03, HLA-DPB1*04:01, HLA-DPB1*06:01	(66)
	AMGIAMSAFAMMFV	3615-3629	II	HLA-DRB1*04:04	(66)
	MLVYCFLGYFCTCYF	3780-3794	II	HLA-DPA1*01:03, HLA-DPB1*06:01	(66)
	EFSSLPSYAAFATAQ	3947-3961	II	HLA-DQA1*02:01, HLA-DQB1*03:01	(66)

	1) VPLNIPLTTAAKLM 2) NIPLTTAAKLMVVI	4057-4071	II	HLA-DQA1*01:02, HLA-DQA1*02:01, HLA-DQB1*03:01, HLA-DQB1*05:01	(66)
	LCTHTGTGQAITVTP	4298-4312	II	HLA-DQA1*02:01, HLA-DQB1*03:01	(66)
	LKGKYVQIPTTCAND	4345-4359	II	HLA-DRB1*01:01	(66)
	NFNVLVSTVFPPTSF	4704-4718	II	HLA-DPA1*01:03, HLA-DPB1*06:01	(66)
	QKLLKSIAATRGATV	4965-4979	II	HLA-DRB1*01:01	(66)
	SHRFYRLANCAQVL	5041-5055	II	HLA-DRB1*01:01	(66)
	TSSGDATTAYANSVF	5072-5086	II	HLA-DQA1*02:01, HLA-DQB1*03:01	(66)
	GLVASIKNFKSVLYY	5166-5180	II	HLA-DRB1*04:04	(66)
	IVYTACSHAAVDALC	5628-5642	II	HLA-DQA1*02:01, HLA-DQB1*03:01	(66)
	RRNVATLQAENVTGL	5918-5932	II	HLA-DRB1*01:01	(66)
ORF3a	<b>PSDFVRATATIPIQASL</b>	25-41	II	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DQA1*02:01, HLA-DQA1*05:01, HLA-DQB1*03:01	(66), En este estudio
	1) LLLLFTVYSHLLL 2) FVCNLLLLFTVYSH	83-97	II	HLA-DPA1*01:03, HLA-DPB1*06:01	(66)
	<b>CIPYNSVTSSIVTSG</b>	157-172	II	HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DQA1*02:01, HLA-DQA1*05:01, HLA-DQB1*03:01	(66), En este estudio
ORF6	1) <b>EILLIIMRTFKVSIW</b> 2) FKVSIWNLDYIINLI 3) KVSIWNLDYIINLII	13-27	II	HLA-DRB1*01:01, HLA-DQA1*01:01, HLA-DQB1*05:01	(66), En este estudio
ORF7a	MKILFLALITLATC	1-15	II	HLA-DPA1*01:03, HLA-DPB1*06:01	(66)
	1) <b>DNKFALTCFSTQFAFACP</b> 2) DNKFALTCFSTQFAF 3) KFALTCFSTQFAFAC	51-68	II	HLA-DPA1*01:03, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*04:01, HLA-DPB1*06:01, HLA-DPB1*23:01	(66), En este estudio
	<b>VKHVYQLRARSVSPKLF</b>	71-88	II	HLA-DRB1*01:01, HLA-DRB1*07:01	(66), En este estudio
	1) VAAIVFITLCFTLKR 2) LIVAAIVFITLCFTL	104-118	II	HLA-DPA1*01:03, HLA-DPB1*06:01	(66)
ORF8	<b>LVFLGIITVAAFHQ</b>	4-18	II	HLA-DQA1*01:02, HLA-DQB1*05:01	(66), En este estudio
	<b>VRCSFYEDFLEYHDV</b>	100-114	II	HLA-DQA1*01:01, HLA-DQB1*05:01	(66), En este estudio
<b>Epitopos promiscuos</b>					
S	<b>QSIIAYTMSLGAENSVAY</b>	690-705	I, II	HLA-A*02:01 HLA-DRB1*04:01, HLA-DRB*07:01	(61)

	<b>LQYGSFCTQLNRALTGI</b>	754-770	I, II	HLA-A*02:01, HLA-DRB1*01:01, HLA-DRB1*04:01	(67), En este estudio
	<b>RSFIEDLLFNK/TLA</b>	815-829	I, II	HLA-A*23:01, HLA-A*11:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01	(67), En este estudio
	<b>GWTFGAGAALQIPFAMQMAY RFNGIGVTQNVLY</b>	885-917	I, II	HLA-B*15:01, HLA-B*40:01, HLA-B*58:01, HLA-C*15:02, HLA-A*11:01, HLA-A*68:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DPA1*01:03, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*03:01, HLA-DPB1*23:01	(67)
	<b>LQDVVNQNAQALNTLVKQLSS NFGAISSVLNDIL</b>	948-981	I, II	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*68:01, HLA-B*07:02, HLA-DRB1*01:01	(61), (65), (67)
	<b>SVLNDILSRDKVEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEIR ASANLAATKMSE</b>	975-1031	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*33:01, HLA-A*68:01, HLA-B*27:05, HLA-B*40:01, HLA-B*40:02, HLA-B*44:02, HLA-B*44:03, HLA-B*45:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DQA1*05:01, HLA-DQB1*03:01	(61), (67)
	<b>WLGFIAGLIAIVMVTI</b>	1217-1232	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-A2, HLA-DRB1*01:01, HLA-DQA1*05:01, HLA-DQB1*03:01	(67)
M	<b>FLWLLWPVTLACFVLAAYRI NWITGGIAIAMACLVGLMWLS YFIASFRLFARTRSMWSFNP</b>	53-114	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*23:01, HLA-A*24:02, HLA-A*29:02, HLA-A*32:01, HLA-A*68:02, HLA-B*07:02, HLA-B*27:05, HLA-A*32:01, HLA-B*35:01, HLA-B*39:01, HLA-B*51:01, HLA-B*53:01, HLA-B*54:01, HLA-B*57:01, HLA-B*58:01, HLA-C*06:02, HLA-C*07:01, HLA-DRB1*01:01, HLA-DRB1*01:02, HLA-DRB1*03:05, HLA-DRB1*03:09,	(61), (66), (67), (58), (69), (70), (71), En este estudio

				HLA-DRB1*04:01, HLA-DRB1*04:05, HLA-DRB1*04:08, HLA-DRB1*04:26, HLA-DRB1*07:01, HLA-DRB1*07:03, HLA-DRB1*08:03, HLA-DRB1*08:13, HLA-DRB1*09:01, HLA-DRB1*11:01, HLA-DRB1*11:02, HLA-DRB1*11:14, HLA-DRB1*11:20, HLA-DRB1*11:21, HLA-DRB1*11:28, HLA-DRB1*13:02, HLA-DRB1*13:05, HLA-DRB1*13:07, HLA-DRB1*13:22, HLA-DRB1*13:23, HLA-DRB1*15:01, HLA-DRB1*15:02, HLA-DRB1*15:06, HLA-DQA1*02:01, HLA-DQA1*03:01, HLA-DQB1*03:01, HLA-DQB1*03:02, HLA-DQB1*03:03, HLA-DQB1*04:01	
	IKDLPKEITVATSRTL <sup>S</sup> YYKLGASQRVAGDSGFAAYSR	161-198	I, II	HLA-A*01:01, HLA-A*03:01, HLA-A*11:01, HLA-A*26:01, HLA-A*29:02, HLA-A*30:02, HLA-A*31:01, HLA-A*33:01, HLA-A*68:01, HLA-B*07:02, HLA-B*35:01, HLA-B*57:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DQA1*02:01, HLA-DQA1*05:01, HLA-DQB1*03:01	(66), (67), (68), En este estudio
N	ATKAYNVTQAFGRRG	264-278	I, II	N/A	(67), (70)
	WPQIAQFAPSASAFFGMSRIG MEVTPSGTWL	301-331	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-B*40:01, HLA-DRB1*01:01, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*08:03, HLA-DQA1*05:01, HLA-DQB1*03:01	(61), (67), (69)
	LLNKHIDAYKTFPPEPKKD	352-371	I, II	HLA-A*03:01, HLA-A*11:01, HLA-A*31:01, HLA-A*68:01, HLA-DRB1*01:01	(67)
E	TGTLIVNSVLLFLAFV <sup>F</sup> LLVTL AILTARLCAYC	9-43	I, II	HLA-A*01:01, HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*03:01, HLA-A*11:01, HLA-A*26:01, HLA-A*29:02, HLA-A*30:02, HLA-A*31:01, HLA-A*33:01,	(66), (67)

				HLA-A*68:01, HLA-A*68:02, HLA-B*40:01, HLA-B*45:01, HLA-DPA1*01:03, HLA-DPB1*02:01, HLA-DPB1*04:02, HLA-DPB1*05:01	
ORF1ab	<b>AIIASFSASTSAFVETVKG</b>	474-493	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-DQA1*01:02, HLA-DQA1*02:01, HLA-DQB1*03:01, HLA-DQB1*05:01	(66), (67)
	<b>NYMPYFFTLQLLCTFT</b>	2167-2183	I, II	HLA-A*24:02, HLA-DPA1*01:03, HLA-DPB1*06:01	(66)
	<b>TAFGLVAEWFLAYILFTRFFYV LGAAI</b>	2319-2346	I, II	HLA-A*02:01, HLA-B*40:01, HLA-B*58:01, HLA-DPA1*01:03, HLA-DPB1*06:01	(66)
	<b>MWLIINLVQMAPISAM</b>	2365-2380	I, II	HLA-B*15:01, HLA-DQA1*01:02, HLA-DQB1*05:01	(66)
	<b>RMYIFFASFYVVKSY</b>	2382-2397	I, II	HLA-A*03:01, HLA-A*11:01, HLA-A*26:01, HLA-A*31:01, HLA-A*33:01, HLA-A*68:01, HLA-A*69:01, HLA-B*15:01, HLA-B*58:01, HLA-DPA1*01:03, HLA-DPB1*06:01	(66), (67)
	<b>FLLNKEMYLKLRSVDLLPL</b>	3183-3201	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-DRB1*01:01	(66)
	<b>PPQTSITSAVLQSGFRKM</b>	3252-3269	I, II	HLA-A*03:01, HLA-A*11:01, HLA-DQA1*02:01, HLA-DQB1*03:01	(66), (67)
	<b>KYKQVRIQPGQTFVLA</b>	3363-3379	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-DRB1*01:01	(66), (67)
	<b>KHAFLCLFLLPSLATVAYFN</b>	3632-3651	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-DRB1*01:01, HLA-DPA1*01:03, HLA-DPB1*06:01	(66), (67)
	<b>YFNMVYMPASWVMRI</b>	3649-3663	I, II	HLA-B*07:02, HLA-B*35:01, HLA-DRB1*01:01	(66), (67)
	<b>SMWALIISVTSNYSQVTTVM</b>	3732-3747	I, II	HLA-A*02:01, HLA-DQA1*01:02, HLA-DQA1*02:01, HLA-DQB1*03:01, HLA-DQB1*05:01	(66)
	<b>PLIVTALRANSVAVKLQ</b>	4125-4140	I, II	HLA-A*03:01, HLA-A*31:01, HLA-A*11:01, HLA-DRB1*01:01, HLA-DQA1*01:02, HLA-DQB1*05:01	(66), (67)
	<b>RQMSCAAGTTQACTDDNAL AY</b>	4150-4171	I, II	HLA-A*01:01, HLA-DQA1*02:01,	(66)



				HLA-DQB1*03:01	
	<b>LLMPILTLTRALTAE</b>	4632-4646	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-DQA1*01:02, HLA-DQB1*05:01	(66), (67)
	<b>MPNMLRIMASLVLAR</b>	5018-5032	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-DQA1*01:02, HLA-DQB1*05:01	(66), (67)
	<b>LMIERFVSLAIDAYPL</b>	5247-5261	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-B*08:01, HLA-DPA1*01:03, HLA-DPB1*03:01	(66), (67)
	<b>PEFYEAMYTPHTVLQAVGACV LCN</b>	5310-5333	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-DRB1*01:01	(66), (67)
	<b>ASKILGLPTQTV DSS</b>	5846-5860	I, II	HLA-A*02:01, HLA-B*15:01, HLA-DRB1*01:01	(66), (67)
	<b>MTYRRLISMMGFKMN</b>	5974-5988	I, II	HLA-A*03:01, HLA-DRB1*01:01, HLA-DPA1*01:03, HLA-DPB1*03:01	(66)
ORF3a	<b>KKRWQLALS KGVHFVCN</b>	66-82	I, II	HLA-B*27:05, HLA-DRB1*01:01	(66)
	<b>EAPFLYLYALVYFLQSINFVRII MR</b>	102-126	I, II	HLA-A*02:01, HLA-DRB1*01:01, HLA-DPA1*01:03, HLA-DPA1*03:01, HLA-DPB1*04:01, HLA-DPB1*04:02, HLA-DPB1*06:01	(66)
	<b>VLHSYFTSDYYQLYSTQLST</b>	202-221	I, II	HLA-A*01:01, HLA-DRB1*01:01, HLA-DPA1*01:03, HLA-DPB1*04:01	(66)
ORF6	<b>MFHLVDFQVTIAEILLII</b>	1-18	I, II	HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*68:02, HLA-B*40:01, HLA-DRB1*01:01, HLA-DRB1*07:01, HLA-DPA1*01:03, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:02, HLA-DPB1*06:01, HLA-DPB1*23:01	(66), (67), En este estudio
ORF7a	<b>CELYHYQECVRGTTVLLKEPC</b>	15-35	I, II	HLA-A*11:01, HLA-A*31:01, HLA-A*68:01, HLA-DRB1*01:01	(67)
ORF7b	<b>IELSLIDFYLCFLAFLFLVLIML IIFWSLELQDHN</b>	2-38	I, II	HLA-A*01:01, HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:06, HLA-A*23:01, HLA-A*24:02, HLA-A*26:01, HLA-A*29:02, HLA-A*68:02, HLA-B*18:01, HLA-B*40:01,	(66), (67), En este estudio

				HLA-B*44:02, HLA-B*44:03, HLA-B*45:01, HLA-DRB1*01:01, HLA-DPA1*01:03, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*06:01 HLA-DPB1*23:01, HLA-DQA1*01:01, HLA-DQB1*05:01	
--	--	--	--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

\* Dentro de las casillas de los epítomos de las clases I y II se desglosan todas las secuencias que han sido descritas por otros autores y en negrita se señalan los epítomos predichos en este estudio que coincidían con los reportados. Para los epítomos promiscuos se señalan en negrita el *core* del péptido.

Por otro lado, se encontraron 17 epítomos de células T reportados experimentalmente y capaces de ser reconocidos para montar una respuesta inmunitaria (Anexo 1, Tabla 9), de los cuales 8 son reconocidos por células T CD4+ (4 de la proteína S: **NLLLQYGSFCTQLNR**, **NFSQILPDPSKPSKR**, **TDEMIAQYTSALLAG**, **QLIRAAEIRASANLAATK**; 2 de la proteína M: **GAVILRGHLRIAGHHLGR**, **TSRTLSYYKLGASQRVA**; 2 de la proteína N: **MKDLSRWYFYLLGTGPEAG**, **DDQIGYYRRATRRIR**). Por otro lado, 6 son reconocidos por células T CD8+ (4 de la proteína S: **YLQPRTFLL**, **ASANLAATK**, **WTAGAAAYY**, **TPCNGVEGF**; 1 de la proteína M: **LLESELVIGAVILRGHLR**; 1 de la proteína N: **KTFPPTEPK**), y 3 son epítomos promiscuos (1 de la proteína S: **CTFEYVSQPFLMDLE**; 2 de la proteína N: **GMEVTPSGTWLTYTGAIKLD**, **LSPRWYFYLLGTGPEAGL**).

A través de la alineación múltiple de secuencia, se identificaron 367 epítomos que compartían cierto porcentaje de identidad entre las proteínas del SARS-CoV-2 y la de los otros coronavirus humanos (SARS-CoV, MERS-CoV, HKU1, NL63, OC43 y 229E). En el Anexo 2, Tabla 11, se muestra el compendio de epítomos, restringidos a MHC-I, MHC-II y promiscuos, con el respectivo porcentaje de identidad entre la secuencia del SARS-CoV-2 y los otros seis coronavirus. Fueron 63 epítomos los que compartieron más del 50% de identidad entre el SARS-CoV-2 y todos los demás coronavirus humanos, 60 de MHC-I, 2 de MHC-II y 1 epítomo promiscuo, siendo los epítomos del SARS-CoV los que tenían los porcentajes de identidad más altos con

el SARS-CoV-2 (Tabla 6). Los porcentajes de identidad para los epítomos de células T obtenidos fueron considerablemente más altos que los encontrados para los epítomos de células B.

**Tabla 6. Epítomos conservados de células T con restricción a MHC-I, MHC-II y promiscuos con identidad superior al 50% entre el SARS-CoV-2 y seis coronavirus humanos.**

	Proteína	SARS-CoV-2	SARS	MERS	229E	HKU1	NL63	OC43
MHC-I	S	AEVQIDRLI	100	66.7	66.7	88.9	66.7	88.9
	ORF1ab	FSAVGNICY	100	100	88.9	88.9	77.8	88.9
		AMRPNFTIK	88.9	77.8	55.6	55.6	66.7	55.6
		LVSTQEFYR	100	77.8	55.6	55.6	55.6	55.6
		TELEPPCRF	100	66.7	88.9	66.7	66.7	66.7
		KVKYLYFIKGL	100	54.5	63.6	72.7	72.7	72.7
		LQAGNATEV	100	66.7	66.7	77.8	66.7	77.8
		VYRAFQDIYNDK	91.7	64.3	50	58.3	58.3	66.7
		AVAKHDFFK	88.9	77.8	88.9	66.7	66.7	66.7
		NWYDFGDFI	88.9	77.8	77.8	88.9	77.8	66.7
		KLFDYFYKYWDQTY	92.9	78.6	64.3	57.1	50	50
		FPPTSFGPLVRK	100	66.7	66.7	66.7	66.7	66.7
		VPFVSTGYHFR	100	75	58.3	75	58.3	75
		TVKPGNFNK	100	77.8	88.9	88.9	66.7	88.9
		FAQDGNAAISDYDY	100	85.7	64.3	78.6	57.1	78.6
		YYRYNLPTM	100	88.9	88.9	88.9	66.7	88.9
		QVIVNNLDK	100	77.8	55.6	88.9	55.6	88.9
		KSAGFPFNK	100	88.9	77.8	88.9	77.8	88.9
		NVIPTITQM	100	88.9	55.6	77.8	66.7	77.8
		SAKNRARTV	100	100	77.8	100	77.8	100
		MTNRQFHQK	100	88.9	88.9	77.8	77.8	77.8
		KFYGGWHNMLK	100	81.8	90.9	72.7	81.8	72.7
		NPHLMGWDPYK	90.9	100	81.8	90.9	90.9	90.9
		CSLSHRFYRL	90	60	60	70	60	80
		RLANCAQVL	100	100	80	100	80	100
		SSGDATTAY	100	100	77.8	88.9	77.8	88.9
		SVFNICQAV	100	77.8	88.9	100	77.8	100
		LSDDAVVCF	88.9	77.8	77.8	77.8	77.8	77.8
		LMIERFVSLAIDAYPLTK	100	94.4	72.2	83.3	72.2	83.3
		DVTDVTQLY	100	66.7	66.7	77.8	55.6	77.8
		NGQVFGLYK	100	88.9	77.8	88.9	66.7	88.9
		RPPLNRNYV	100	88.9	77.8	77.8	77.8	77.8
		AAVDALCEK	100	100	77.8	100	77.8	100
		KVQIGEYTFEK	100	72.7	63.6	54.5	72.7	54.5
		IPARARVEC	100	88.9	100	66.7	100	66.7
		ALPETTADI	100	88.9	77.8	66.7	77.8	66.7
		IVVFDEISM	100	66.7	77.8	77.8	77.8	77.8
		VYIGDPAQL	100	88.9	77.8	100	77.8	100
		LPAPRTLLT	100	100	66.7	77.8	66.7	77.8
		CPAEIVDTV	100	77.8	88.9	77.8	88.9	88.9
EAVGTNLPL	88.9	66.7	66.7	66.7	55.6	55.6		
VVRIKIVQM	100	66.7	66.7	77.8	66.7	77.8		
NLSDRVVFLW	90.9	54.5	63.6	63.6	81.8	63.6		
HSIGFDYVY	88.9	77.8	66.7	66.7	66.7	55.6		
FMDVQQWVG	100	77.8	66.7	55.6	55.6	55.6		
LTRNPAWRKAVF	100	75	50	58.3	66.7	66.7		
SEYDYVIFT	100	77.8	88.9	77.8	77.8	77.8		
VLHDIGNPK	100	66.7	77.8	66.7	88.9	66.7		

		KFTDGVCLFW	100	70	60	60	60	60
		KQLPFFYYS	100	88.9	66.7	77.8	55.6	77.8
		HANEYRLYL	88.9	66.7	66.7	66.7	55.6	77.8
		SLENVAFNVVVK	91.7	66.7	83.3	66.7	83.3	58.3
		SQLGGLHLI	80	80	80	70	80	70
		TTLPVNVAFELWAKRNI	100	70.6	64.7	52.9	76.5	64.7
		PYNMRVIHF	100	88.9	88.9	88.9	77.8	88.9
		DSKEGFFTY	100	66.7	88.9	77.8	77.8	77.8
		ALGGSVAIK	77.8	100	77.8	88.9	88.9	100
		KITEHSWNADLYK	100	69.2	61.5	92.3	69.2	84.6
		EAFLLGCNYLG	90.9	72.7	72.7	81.8	90.9	72.7
		YVMHANYIFWRNTNPI	93.8	68.8	50	62.5	56.2	62.5
MHC-II	S	QLSSNFGAISSVLN	100	50	57.1	64.3	50	64.3
		EVQIDRLITGRLQS	100	64.3	64.3	71.4	64.3	71.4
MHC-I & MHC-II	S	RSFIEDLLFNKVTLA	100	80	53.3	73.3	60	73.3

\*La tabla está coloreada según el porcentaje de conservación (más oscuro cuanto mayor porcentaje de identidad)

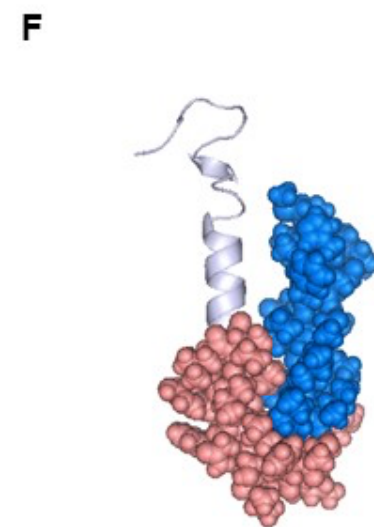
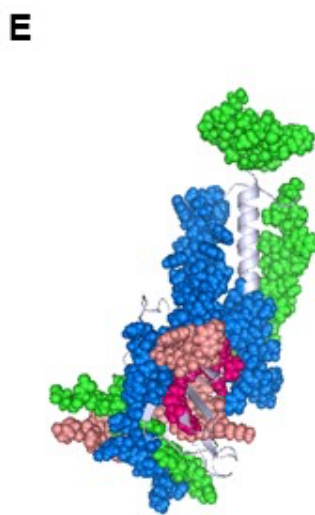
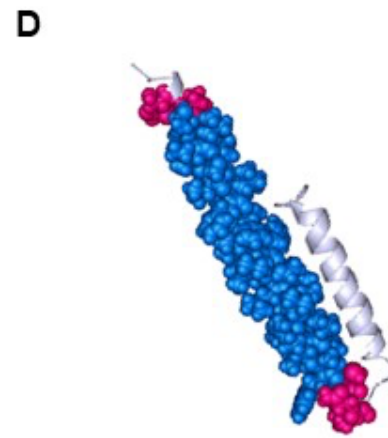
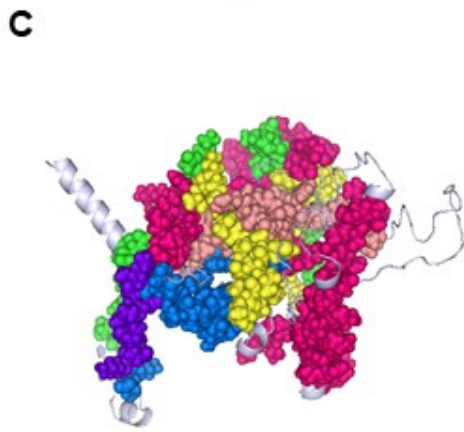
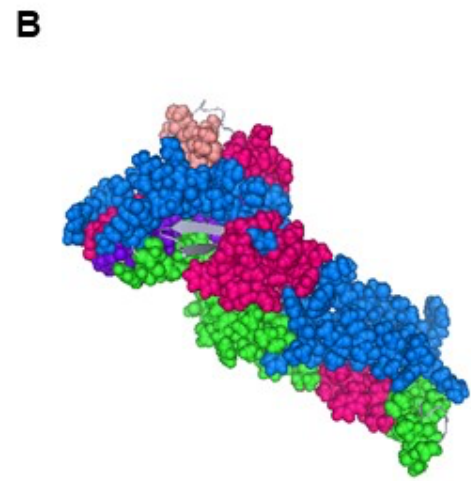
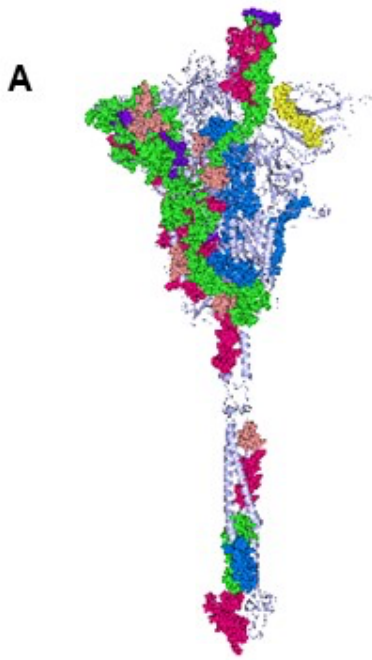
Para la visualización de los epítomos se utilizó un modelo 3D optimizado tanto para la clase I y la clase II, en el cual se representaron los epítomos reportados en la literatura previamente (en color magenta), los epítomos predichos en este estudio (en color rosa), así como los epítomos reportados experimentalmente y capaces de ser reconocidos para montar una respuesta inmunitaria por células T CD4+ y CD8+ (en color morado).

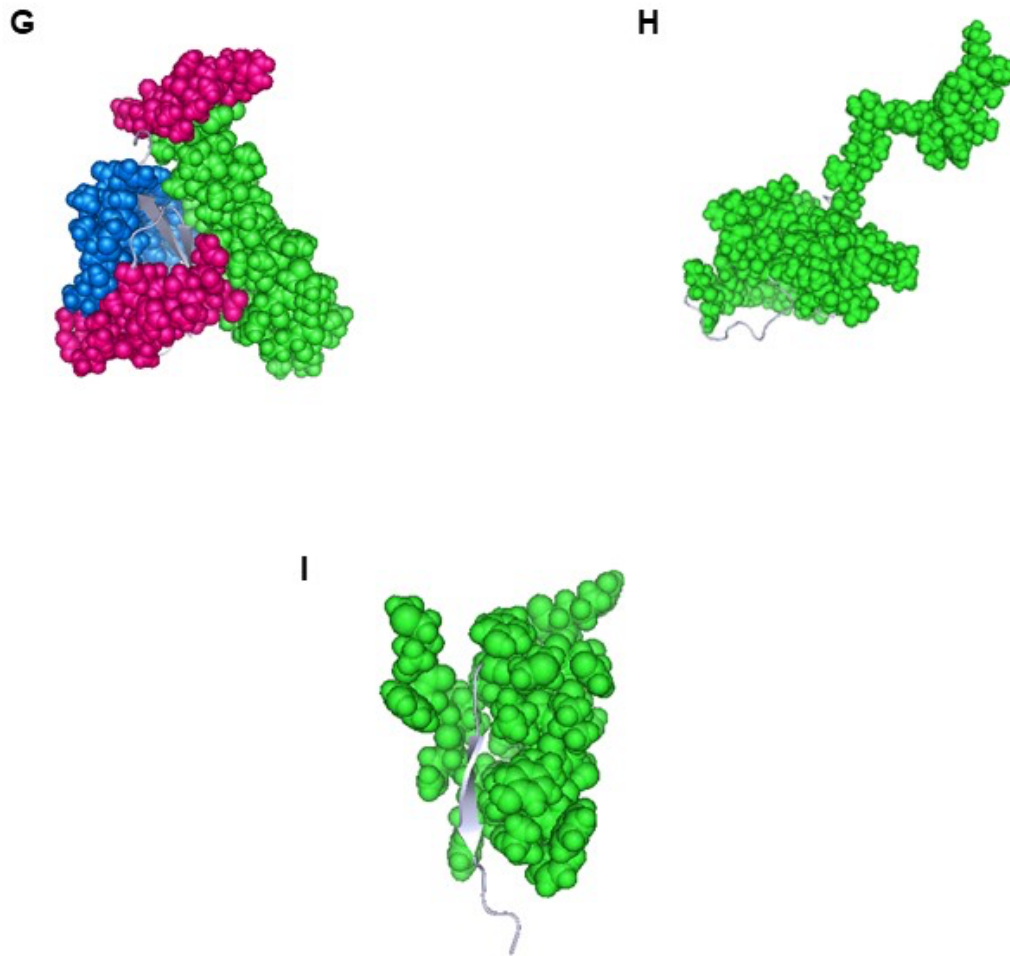
Del mismo modo, se representaron los epítomos promiscuos reportados (en color azul), los epítomos promiscuos predichos en este estudio (en color verde) y los epítomos promiscuos reportados experimentalmente (en color amarillo).

Dichos modelos mostraron que los epítomos descritos en la bibliografía se pueden encontrar en toda la estructura de la glicoproteína de espícula. Por el contrario, los epítomos predichos en este estudio están presentes en toda la estructura con la excepción de una brecha notable en el dominio TM. Los epítomos promiscuos predichos y reportados, así como los epítomos analizados experimentalmente se localizaron en las regiones que comprenden la subunidad S1, a diferencia de los epítomos promiscuos ya evidenciados experimentalmente que se encontraron preferentemente en el dominio NTD (Figura 10-A; Figura 11-A). En el caso de la proteína M, se logró evidenciar que los epítomos tanto reportados como los predichos, se distribuyen por completo en toda la proteína, sin embargo, los epítomos de los cuales ya hay evidencia experimental están principalmente

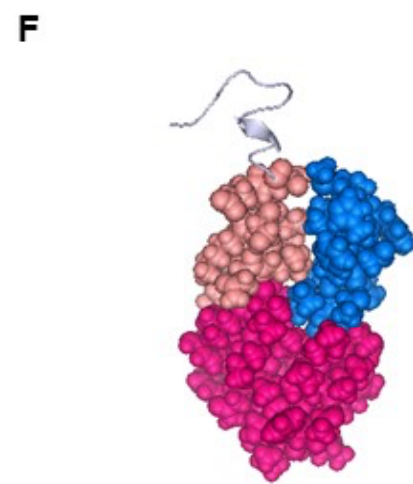
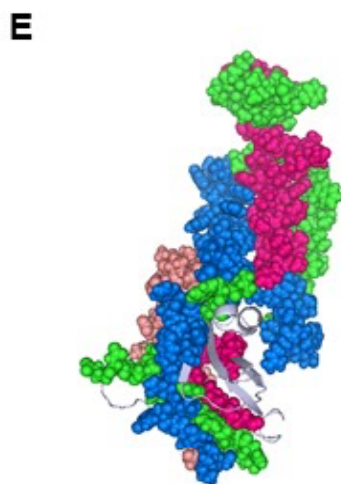
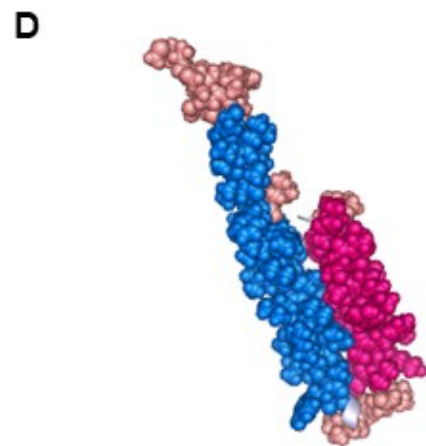
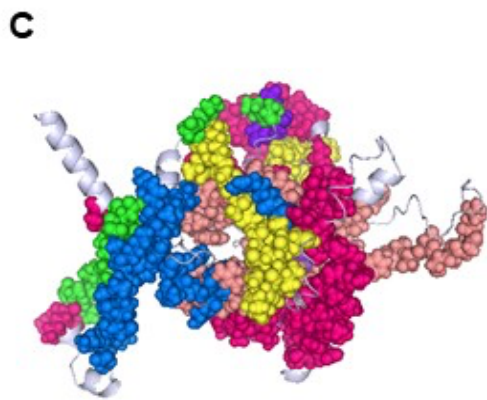
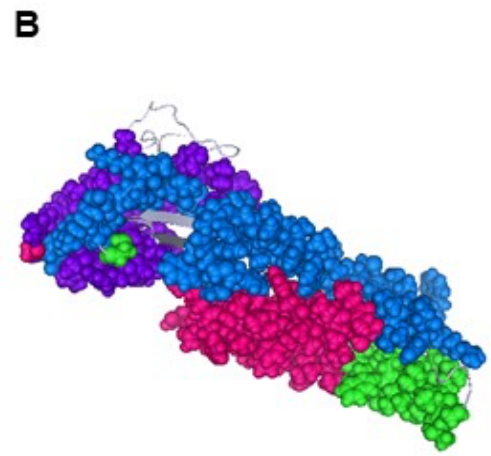
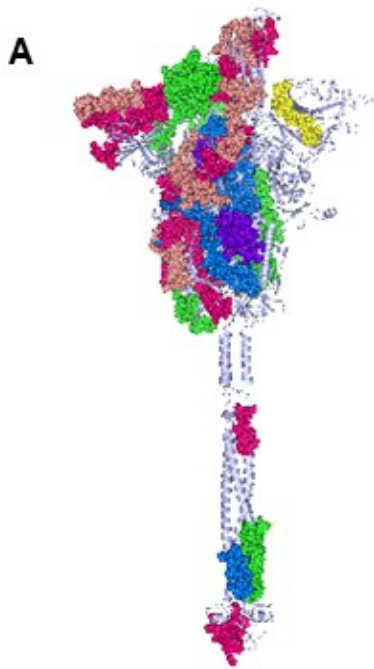
centralizados en el dominio N-terminal. De igual forma todos los epítomos promiscuos se observaron a lo largo de toda la proteína (Figura 10-B; Figura 11-B). Por otro lado, al observar los epítomos de la proteína N en el modelo 3D, se sugiere que la distribución de éstos es en toda la proteína, destacando que aquellos que ya tienen un respaldo experimental fueron mapeados principalmente en el dominio de unión al RNA, el sitio de unión y en el dominio de dimerización (Figura 10-C; Figura 11-C), muy similar a lo que se observa con los epítomos B lineales en esta misma. Para la proteína E todos los epítomos con restricción a MHC-I se situaron en el dominio N-terminal (Figura 10-D), mientras que los epítomos restringidos a MHC-II se encontraron a lo largo de toda la proteína (Figura 11-D).

Con respecto a los epítomos de la proteína ORF3a éstos se encontraron a lo largo de las tres hélices transmembranales (Figura 10-E; Figura 11-E), lo mismo para los de la proteína ORF6, los cuales se distribuyeron en ambas  $\alpha$  hélices de la proteína (Figura 10-F; Figura 11-F). Los epítomos de la proteína ORF7a se localizaron tanto en el dominio N-terminal como en el dominio C-terminal (Figura 10-G; Figura 11-G), al igual que la proteína ORF8 cuyos epítomos se situaron a lo largo del dominio N-terminal (Figura 10-H; Figura 11-H). Finalmente, los epítomos de la proteína ORF10 con restricción a MHC-I y MHC-II se localizaron en toda la proteína, con la distinción de que los epítomos predichos para la clase II se observaron en su mayoría en una de las láminas  $\beta$  plegadas.

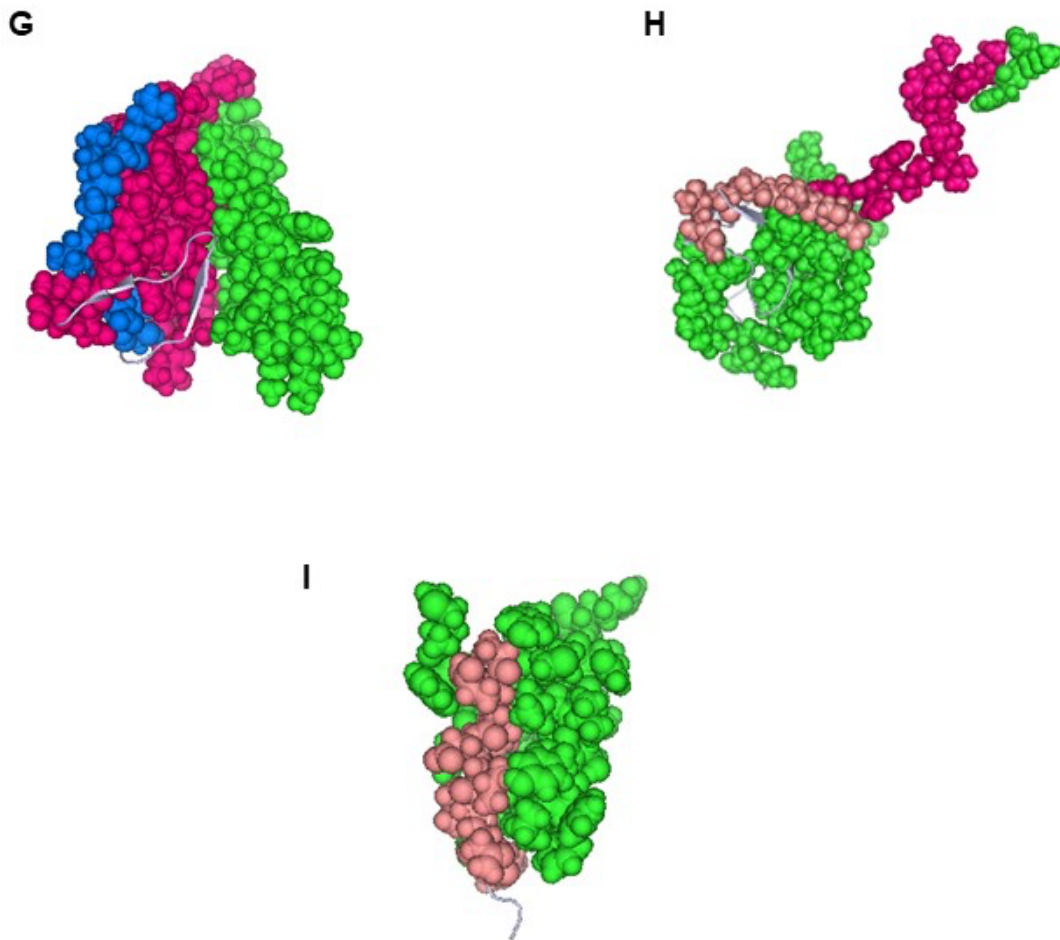




**Figura 10. Epítomos de células T con restricción a MHC-I y promiscuos de las proteínas S (A), M (B), N (C), E (D), ORF3a (E), ORF6 (F), ORF7a (G), ORF8 (H) y ORF10 (I).** En magenta se representan los epítomos reportados en la literatura, en rosa los epítomos que se obtuvieron de las predicciones en este estudio, en morado los epítomos analizados experimentalmente y capaces de ser reconocidos para montar una respuesta inmunitaria, en azul los epítomos promiscuos reportados, en verde los epítomos promiscuos predichos en este estudio y en amarillo los epítomos promiscuos reportados experimentalmente.







**Figura 11. Epítomos de células T con restricción a MHC-II y promiscuos de las proteínas S (A), M (B), N (C), E (D), ORF3a (E), ORF6 (F), ORF7a (G), ORF8 (H) y ORF10 (I).** En magenta se representan los epítomos reportados en la literatura, en rosa los epítomos que se obtuvieron de las predicciones en este estudio, en morado los epítomos analizados experimentalmente y capaces de ser reconocidos para montar una respuesta inmunitaria, en azul los epítomos promiscuos reportados, en verde los epítomos promiscuos predichos en este estudio y en amarillo los epítomos promiscuos reportados experimentalmente.

## DISCUSIÓN

Hoy, todavía estamos en la batalla para detener la pandemia causada por el virus SARS-CoV-2 la cual se ha posicionado como un problema de salud pública prioritario, ya que desde diciembre de 2019 y hasta principios de 2022 ha provocado la muerte de más de seis millones de personas a nivel mundial<sup>2</sup>. Por si fuera poco, el surgimiento y expansión de nuevas variantes virales por todo el mundo, ha puesto en riesgo los esfuerzos actuales de control de la pandemia. Lo anterior, a consecuencia de que el genoma del SARS-CoV-2 ha sufrido mutaciones afectando principalmente al gen que codifica para la proteína de la espícula (S). Las mutaciones en este antígeno se han establecido en regiones específicas y la importancia de esta proteína debe ser enfatizada ya que interactúa con la enzima convertidora de angiotensina 2 (ACE2), el receptor específico de la célula huésped, dando paso a una selección de variantes con cambios beneficiosos para la cepa original en términos de su capacidad infecciosa, transmisibilidad y resistencia a anticuerpos neutralizantes<sup>88</sup>. Adicionalmente, debido a que dicha proteína está expuesta en la superficie del virus y media la entrada en las células huésped, es el objetivo principal de los anticuerpos neutralizantes tras la infección y el foco del diseño terapéutico y de las vacunas<sup>6</sup>.

Por lo anterior, es evidente que aún carecemos de una comprensión profunda del desarrollo de las vacunas debido a que se están llevando a cabo sin un conocimiento completo de la patogénesis del virus, la respuesta del huésped a la infección, las respuestas inmunitarias potencialmente perjudiciales y la inmunidad protectora. Además, la mayoría de los estudios clínicos realizados se han centrado principalmente en la producción de anticuerpos neutralizantes, en donde si bien se ha observado que muchas de las personas presentan estos anticuerpos después de la vacunación, también se ha hecho evidente que en algunos casos esta respuesta no tiene larga duración. Por consiguiente, es que resulta importante prestar mayor atención a la inmunidad celular, de la cual se ha sugerido podría desempeñar un papel fundamental para combatir a este nuevo coronavirus.

A través de la revisión de la literatura y los ejercicios de predicción de epítomos llevados a cabo en este proyecto, se recolectaron 235 epítomos lineales

para células B de las proteínas de SARS-CoV-2, reportando 171 epítomos nuevos, la mayoría de ellos situados en las proteínas S (RBD, subunidades S1/S2 y sitio de furina) y ORF1ab, impulsados por sus grandes tamaños dentro del proteoma del SARS-CoV-2. Estudios recientes han informado la identificación de anticuerpos dirigidos a estos sitios particulares, por lo tanto, consideramos que los nuevos epítomos en este estudio deben investigarse como posibles blancos protectores<sup>72,75</sup>.

Además, se ha confirmado la actividad neutralizante de anticuerpos contra el antígeno S, principalmente en el RBD<sup>89</sup>. De hecho, los anticuerpos dirigidos contra el SARS-CoV, denominados CR3022 y 47D11, pueden reaccionar de forma cruzada con el RBD del SARS-CoV-2. Por lo tanto, el RBD de la proteína S representa un objetivo importante para la neutralización mediada por anticuerpos, el desarrollo de vacunas y otras estrategias antivirales<sup>25</sup>.

Por otro lado, se ha demostrado que el epítomo *GDEVQRQIAPGQTGKIADYNYKLPDD*, el cual se traslapa con uno de los epítomos predichos en este trabajo (Tabla 7, Anexo 1), puede provocar una respuesta de anticuerpos neutralizantes en ratón<sup>90</sup>. De igual forma, diversos trabajos han demostrado la actividad inhibitoria y neutralizante de diversos anticuerpos monoclonales dirigidos a este sitio, dado que interaccionan con múltiples residuos en esta proteína<sup>89,91</sup>.

Así como el epítomo mencionado anteriormente, el epítomo *ASYQTQTNSPRRARSVASQ* (Tabla 7, Anexo 1), predicho en este estudio, también ha sido descrito como *YQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSN* por Polyiam y colaboradores<sup>75</sup>. Ellos señalan que dicho epítomo se encuentra en el sitio de escisión de la S1/S2, donde la secuencia *RRAR*, es reconocida y escindida por la proteasa furina, lo que da como resultado la separación de los dominios S1 y S2 durante el ensamblaje del virus. De generarse anticuerpos dirigidos a esta región inmunodominante de la proteína, éstos podrían bloquear el procesamiento de escisión de la proteína S durante la colonización del virus a la célula huésped<sup>75</sup>.

Con respecto a los epítomos *ILPDPSKPSKRS* y *FIEDLLFNKVTADAGFFIKQYGDCLG*, que se encuentran en el sitio de escisión de S2 (residuos 805-842), y el epítomo *PSKPSKRSFIEDLLFNKV*, son epítomos

neutralizantes<sup>92</sup> y, en consecuencia, los anticuerpos que se unen a estas regiones dan como resultado la inhibición de la escisión.

Por otro lado, como parte de la recopilación bibliográfica de epítomos, se localizó un epítomo en el sitio citoplasmático de la proteína S (*CKFDEDDSEPVLKGVKLHYT*) el cual se ha identificado como un epítomo inmunodominante<sup>71,93</sup>. Aunque no se ha demostrado que este epítomo genere anticuerpos neutralizantes contra el virus SARS-CoV-2, una investigación con una variante de coronavirus sugiere su potencial, pues se ha demostrado que epítomos que ya fueron comprobados experimentalmente de células B ubicados en el sitio citoplasmático de la proteína S del virus de la diarrea epidémica porcina (PEDV), un miembro de los alfa coronavirus, son un objetivo importante de anticuerpos neutralizantes<sup>94</sup>.

La proteína S, no ha sido el único antígeno que se ha utilizado para la identificación de epítomos inmunodominantes en el SARS-CoV-2. También han sido estudiadas las proteínas M y N, de las cuales ya se ha expuesto, mediante ensayos serológicos, que son inmunodominantes entre los antígenos del coronavirus y por ende también pueden ser óptimos candidatos para ayudar a controlar la transmisión del virus<sup>76</sup>; de hecho, se ha descrito que la proteína M induce una fuerte respuesta humoral y los epítomos antigénicos se han identificado en las regiones TM1 y TM2 de dicha proteína. Por tanto, la proteína M puede ser un inmunógeno potencial en aplicaciones terapéuticas<sup>25</sup>.

Por otra parte, existen reportes que han descrito que niveles altos de IgG, dirigidos a los epítomos para células B lineales de la proteína S como *TESNKKFLPFQQFGRDIA* y *PSKPSKRSEFIEDLLFNK*, así también como el epítomo *NNAAIVLQLPQGTTLPKG* de la proteína N, están asociados a la gravedad de la enfermedad; incluso se ha detectado que los pacientes que desarrollan neumonía e ingresan a la Unidad de Cuidados Intensivos (UCI) mantienen niveles altos de anticuerpos dirigidos a estos epítomos. En el caso del epítomo de la proteína N, se ha relacionado con la linfopenia en pacientes con COVID-19. Estos resultados han sugerido que anticuerpos dirigidos a estos epítomos pueden ser marcadores del desarrollo de la inmunopatología causada por SARS-CoV-2, pues, además, cabe

resaltar, que estos epítomos poseen una baja tasa de mutación, por debajo del 2%<sup>95</sup>. En este estudio, se logró predecir porciones parciales de estos tres epítomos, con lo que se sugiere que es importante realizar ensayos de neutralización e inhibición con estos péptidos para verificar la actividad biológica de los anticuerpos generados a partir de su reconocimiento.

Con respecto a los coronavirus recombinantes que carecen de proteína E se ha visto que presentan títulos virales reducidos, maduración viral alterada y propagación viral incompetente y, por lo tanto, se han sugerido como buenos candidatos a vacunas<sup>25</sup>. De igual forma, se ha descrito que la supresión de la expresión de ORF3a en el SARS-CoV dio como resultado una disminución en la liberación del virus y una menor morbilidad. Además, los canales iónicos son importantes dianas terapéuticas y ya se han desarrollado muchos fármacos de canales iónicos. Por tanto, ORF3a puede ser otro potente objetivo farmacológico para el tratamiento de enfermedades debido a su papel en la patogenicidad<sup>25</sup>.

Por lo que se refiere a los epítomos de células T, en este estudio se recopilaron un total de 245 posibles epítomos de células T, de los cuales 120 han predicho la unión a MHC-I, 87 han predicho la unión a MHC-II y 38 fueron epítomos promiscuos, todos los cuales han sido reportados en la literatura. En tanto que los epítomos predichos en este estudio sumaron un total de 391 posibles epítomos de células T, de los cuales 327 han predicho la unión a MHC-I, 30 han predicho la unión a MHC-II y 34 se pronostica que son epítomos promiscuos de células T. Aunque la mayoría de las vacunas actualmente desarrolladas se enfocan en la generación de anticuerpos neutralizantes con la premisa de protección, estudios recientes sugieren el fuerte papel que juegan las células T en la protección contra la infección.

La infección por SARS-CoV-2 puede desencadenar una fuerte inmunidad de células T en los pacientes, y la primera autopsia de pacientes con SAR-CoV-2 mostró la acumulación de células T en el pulmón. Con la infección por SARS-CoV-2, cada vez hay más pruebas de que se producen células T específicas del virus. De igual manera, se ha señalado que el HLA juega un papel importante en la respuesta inmune a la infección viral, ya que está asociado con una mayor gravedad de la infección por SARS-CoV-2. Por lo tanto, algunos investigadores han analizado

la afinidad de unión entre HLA y los péptidos de SARS-CoV-2 a través de la predicción bioinformática, para la obtención de muchos HLA ventajosos. Otros experimentos han encontrado que los genotipos HLA Clase I juegan un papel importante en la infección y progresión del SARS-CoV-2<sup>96</sup>.

Un estudio en pacientes con COVID-19 leve y grave mostró que, como resultado de la infección, los individuos desarrollaron células T específicas del SARS-CoV-2 de memoria central y fenotipo de memoria efectora, pero lo que es más importante, es que los casos leves generaron frecuencias más altas de células T CD8+ productoras de múltiples citocinas<sup>97</sup>. También se han detectado fuertes respuestas de células T de memoria específicas al SARS-CoV-2 en muestras de individuos que tenían infecciones leves o incluso asintomáticas, curiosamente algunas en ausencia de respuestas de anticuerpos<sup>98</sup>. Un estudio de Grifoni y colaboradores informó que las células T responden a la estimulación con péptidos del SARS-CoV-2 en muestras recolectadas antes de que comenzara la pandemia actual, lo que sugiere que las infecciones con coronavirus estacionales provocan respuestas de células T de memoria que reaccionan de forma cruzada con el SARS-CoV-2<sup>61</sup>.

En otro estudio, se analizó la inducción de anticuerpos neutralizantes contra el sarbecovirus cruzado en sobrevivientes de la infección por SARS-CoV que recibieron la vacuna BNT162b2 (Pfizer-BioNTech) contra el SARS-CoV-2, encontrando altos niveles de anticuerpos neutralizantes para SARS-CoV y SARS-CoV-2 después de recibir dos dosis de la vacuna. Asimismo, mostraron la inducción eficiente de anticuerpos neutralizantes y de amplio espectro contra todas las variantes de preocupación, (B.1.1.7 (Alfa), B.1.351 (Beta) y B.1.617.2 (Delta) y cinco sarbecovirus pre-emergentes (coronavirus de murciélago RaTG131, WIV1 y RsSHC014; coronavirus de pangolín GD-1 y GX-P5L), señalando la viabilidad de lograr la neutralización de los sarbecovirus a través del refuerzo entre sus descendientes<sup>99</sup>. Otro estudio, mostró que 22 individuos infectados con COVID-19 tuvieron una fuerte inducción de células T secretoras tanto de IFN $\gamma$  como IL-2 que fueron específicas para epítomos altamente conservados, confirmando la reactividad cruzada de estos epítomos con SARS-CoV-2<sup>100</sup>.

La inmunidad preexistente de células T con reactividad cruzada ha sido un tema de debate, sobre todo en cuanto al umbral de inmunidad colectiva, es decir, esto se refiere a la fracción de la población que necesita ser inmune para prevenir una propagación epidémica en curso de una infección. Sin embargo, en este caso, al tratarse de una nueva infección para la que no existía inmunidad preexistente, la inmunidad colectiva se puede generar mediante la infección con el patógeno o mediante la vacunación<sup>101</sup>. Pese a lo anterior, aún falta por conocer si los seres humanos son capaces de generar inmunidad colectiva, aunque se espera que eventualmente se logre.

La inmunidad de las células T preexistentes al SARS-CoV-2 podría ser relevante porque podría influir en la gravedad de la enfermedad por COVID-19. Es plausible que las personas con un alto nivel de células T CD4+ de memoria preexistente que reconocen el SARS-CoV-2 puedan montar una respuesta inmune más rápida y fuerte al exponerse al SARS-CoV-2 y así limitar la gravedad de la enfermedad. Las células T CD4+ potencialmente podrían facilitar una mayor y más rápida respuesta de anticuerpos neutralizantes contra el SARS-CoV-2. Las células T de memoria CD4+ y CD8+ también podrían facilitar la inmunidad antiviral directa en los pulmones y la nasofaringe de manera temprana después de la exposición<sup>102</sup>. A la espera de la confirmación experimental, este trabajo proporciona un panel de 63 epítomos de células T que podrían ser responsables de estas respuestas de reactividad cruzada.

Se sabe que tanto la infección por SARS-CoV-2 como las vacunas de ARNm provocan una respuesta inmunitaria íntegra tanto de células B como de células T, mientras que la aparición de células T activadas precede la recuperación de COVID-19. Algunos estudios *in vitro* caracterizaron en personas convalecientes, linfocitos T CD4+ y CD8+ que fueron reactivos frente a péptidos de la proteína S<sup>103</sup>. Además, en otro estudio se analizaron muestras de sangre periférica de individuos vacunados con BNT162b2 (Pfizer/BioNTech), de los cuales antes de la vacunación, se tomaron muestras de sangre y se observó que poseían células T específicas hacia la proteína S, lo que, probablemente, indicaría la presencia de reactividad cruzada con otros coronavirus estacionales. La presencia de células T de reactividad cruzada

contra SARS-CoV-2 se asocia con una mejor respuesta humoral y celular a la vacunación, y ya que las células T de reactividad cruzada se dirigen principalmente a la región S2 de la proteína S, es posible que medir anticuerpos anti-RBD no genere información robusta sobre la respuesta humoral que induce la vacunación<sup>104</sup>.

Lo anterior refuerza la evidencia de que esta región es menos polimórfica en comparación con la región S1, por ende, los epítomos presentes en S2 están mayormente conservados entre diferentes coronavirus (Tablas 10 y 11. Anexo 2), sugiriendo que los estudios para la nueva generación de vacunas continúen con la adición de nuevos epítomos potenciales de esta región para así poder tener una vacuna de amplio espectro que no sólo sea capaz de proteger contra SARS-CoV-2, sino también contra otros coronavirus estacionales. En el mismo estudio, se identificaron subpoblaciones de células T CD4+ con una polarización a un fenotipo Th1 y CD8+ policlonales, con una respuesta persistente durante 6 meses tanto en su función como en su proliferación. Así mismo se detectaron células T CD4+ y CD8+ de memoria central inducidas por la vacunación, que también se mantuvieron estables durante 6 meses. La detección de estas poblaciones celulares sugiere inmunidad celular de larga duración contra SARS-CoV-2<sup>104</sup>.

En nuestro estudio pudimos detectar varios epítomos de células T CD4+ y CD8+, así como epítomos promiscuos, de los cuales, la mayoría se encuentran tanto en la región S1 y S2, respaldando lo antes descrito, es decir, que el estudio de epítomos de esta región es imperante para generar información acerca de la posible respuesta inmune de memoria.

Adicionalmente, el uso de la bioinformática para identificar epítomos de células B y T para las proteínas del SARS-CoV-2 pueden usarse como objetivos potenciales para desarrollar preparaciones de vacunas contra el SARS-CoV-2. En un estudio se predijeron epítomos de células B y T de la proteína S de los cuales se seleccionaron diez epítomos de células B y veinte epítomos de células T de acuerdo con su antigenicidad. Para probar su eficacia se realizaron análisis de las características tales como digestión, mutación, toxicidad, sensibilización, humedad y propiedades fisicoquímicas. Los resultados mostraron que el epítomo *KIADYNYKL*, predicho de células T (MHC-I), es un objetivo ideal para la vacuna



peptídica porque no sólo tiene una alta antigenicidad, sino que también es un posible antígeno de superficie sin toxicidad ni mutagenicidad<sup>96</sup>. La vacuna peptídica es una combinación de epítomos de células B y células T, que puede inducir inmunidad humoral y estar mediada por células. Este epítomo *KIADYNYKL* se superpone con el epítomo de células B predicho en este estudio *GDEVQRQIAPGQTGKIADYNYKLPDD*, lo que hace posible participar en la activación inmunitaria de las células B y T al mismo tiempo. Por lo anterior, en este mismo estudio, eligieron dicho epítomo para llevar a cabo el acoplamiento molecular con diferentes moléculas de MHC-I, sus resultados indican que después de que el péptido *KIADYNYKL* se une a las moléculas MHC-I, es posible ser presentado al receptor de células T y activar la inmunidad celular<sup>96</sup>. Otro estudio, predijo los epítomos *YQPYRVVVL* y *RAAEIRASASANLAATK*, los cuales también fueron predichos en este estudio, y que se han identificado como péptidos prometedores para inducir la liberación de IFN $\gamma$ <sup>40</sup>, analizados por medio del servidor web IFNepitope.

Como se mencionó anteriormente, la proteína S, no ha sido el único blanco descrito, pues también se han propuesto las proteínas M y N. En este estudio recopilamos 19 epítomos para M y 31 para la proteína N. Para la proteína M se predijeron 10 epítomos para MHC-I, 5 para MHC-II y 4 epítomos promiscuos, y en el caso de la proteína N se predijeron 18 epítomos para MHC-I, 6 para MHC-II y 7 promiscuos. Además, en la proteína M, los epítomos, tanto reportados como predichos, se distribuyen a lo largo de toda la proteína, sin embargo, los que ya cuentan con evidencia experimental se encuentran localizados en el dominio N-terminal. En la proteína N, estos epítomos se encuentran en el dominio de unión al ARN y el dominio de dimerización. En un estudio con 34 pacientes con COVID-19 grave y leve, 11 de 34 pacientes presentaron respuesta al péptido *GAVILRLRGHLRIAGHHLGR*, 16 de 34 pacientes respondieron al péptido *TSRTLSEYKLGASQRVA* y 3 de 34 al péptido *LLESELVIGAVILRGHLR* de la proteína M, los cuales se identificaron en este estudio; los dos primeros generaron una respuesta de células T CD4+ y el último de células T CD8+<sup>61</sup>. Adicionalmente el epítomo *LRGHLRIAGHHLGRCDIKDL* de la proteína M, se ha descrito como un

epítopo altamente conservado, pues en un estudio realizado por Heide y colaboradores<sup>105</sup>, se observó que este péptido fue reconocido por 12 pacientes de 34 y el cual se polariza a un fenotipo de células T CD4+ de memoria efectora. En este contexto, la información observada sugiere la posible participación de epítopos de la proteína M en la respuesta protectora, lo cual también es relevante para la generación de vacunas que puedan generar mayor protección.

En cuanto a la proteína N, se ha probado experimentalmente como un inmunógeno robusto para inducir la respuesta de células T citotóxicas contra el SARS-CoV y MERS-CoV. Asimismo, es la proteína predominante expresada en el virión durante la etapa temprana de la infección y se ha detectado en la mayoría de los pacientes infectados con SARS-CoV desde el primer día de la infección, por lo que se sugiere que una vacuna basada en esta proteína pueda provocar una respuesta inmunitaria dependiente de células T en el momento oportuno<sup>41</sup>.

Aunado a lo anterior, se ha demostrado que células de sangre periférica de 37 donantes que no habían estado expuestos al SARS-CoV ni al SARS-CoV-2 presentaron células T CD4+ productoras de IFN $\gamma$  específicas para SARS-CoV-2 en 19 donantes procedentes de la proteína N. De igual forma se observó reconocimiento a otras proteínas no estructurales como nsp7 y nsp13, en individuos no expuestos previamente a SARS y SARS-CoV-2. En caso contrario, las muestras de donantes recuperados de COVID-19 y SARS reaccionaron preferentemente a péptidos provenientes de la proteína N<sup>77</sup>. Posteriormente, se realizó la caracterización de la respuesta específica de la proteína N en un donante sin exposición previa al SARS y SARS-CoV-2, en el cual se identificaron células T CD4+ que eran reactivas al epítopo *MKDLSPRWFYLLGTGPEAG* de esta proteína, cuyo epítopo también fue reconocido por pacientes recuperados de COVID-19 y SARS. Cabe resaltar que, este epítopo ya había sido localizado por el grupo de Peng y colaboradores<sup>78</sup>, en 12 pacientes recuperados leves y graves de COVID-19, donde lo clasifican como un epítopo promiscuo e inmunodominante para la proteína N. Una porción de este epítopo fue parte de las predicciones aquí presentadas, el cual se encontró de la posición 101-120, por lo que consideramos que este epítopo es de relevancia, ya que se encuentra en una región con un alto grado de homología con

MERS-CoV, OC43 y HKU1, y la persistencia del reconocimiento de este péptido se ha observado durante al menos un año.

Por todo lo anterior, es de suma importancia dar a conocer aquellos epítodos que podrían ayudar al diseño de la vacuna, así como la evaluación de su inmunogenicidad. De este modo se encontró que las proteínas M y N representan, cada una, entre el 11% al 27% de la respuesta total de las células T CD4+, con respuestas adicionales dirigidas a nsp3, nsp4, ORF3a y ORF8, entre otras. Mientras que para las células T CD8+ se reconocieron las proteínas S y M, con al menos ocho ORFs del SARS-CoV-2 como objetivo<sup>61</sup>.

La información inmunológica que surge día con día indica que si una vacuna candidata es capaz de inducir la producción de células T y anticuerpos neutralizantes probablemente tendrá mayor capacidad de protección ante el virus a largo plazo<sup>61</sup>; sin embargo, la aparición de nuevas variantes se presenta como un nuevo desafío en cuanto a los cambios genéticos del virus y el impacto de su contagiosidad. En la actualidad, la OMS reconoce cinco variantes de preocupación (Alfa, Beta, Gama, Delta y la más reciente Omicron) y siete variantes de interés (Épsilon, Eta, Iota, Kappa, Lambda, Theta y Zeta) de SARS-CoV-2<sup>106</sup> y es probable que en el futuro este número aumente, por lo que los científicos de todo el mundo han lanzado varias investigaciones sobre la eficacia de las vacunas que se administran a la población mundial hoy en día.

Estas variantes han aumentado importantemente el número de infecciones alrededor del mundo, incluso, han colapsado el sector de la salud en algunos países<sup>107</sup>. Igualmente, se ha descrito para algunas de estas variantes un aumento de aproximadamente 20 veces en la afinidad para la ACE2, en comparación con el RDB de la variante de Wuhan, lo que las hace más transmisibles<sup>108</sup> y capaces de escapar de la neutralización de anticuerpos monoclonales debido a la mutación E484K<sup>109</sup>. Otras mutaciones en las distintas variantes provocan una mayor replicación de estas, lo que conduce a mayores cargas virales y mayor transmisión<sup>110</sup>. No obstante, y gracias a la vacunación, se han observado altos niveles de neutralización y protección contra los cuadros graves<sup>111,112</sup> y contra la enfermedad sintomática después de recibir dos dosis<sup>110</sup>. Cabe recalcar que los

anticuerpos neutralizantes no son la única línea de defensa generada a partir de la vacunación, el papel de la inmunidad celular en la protección, aunque aún no está claro, cada vez toma más relevancia. Como se muestra en este y otros estudios, la carga antigénica de las proteínas del virus en términos de epítomos de células T es significativa, esto sugiere que dicha inmunidad puede tener una contribución clave en la protección contra la infección, mismo que debe ser investigado más a fondo.

Debido a que los coronavirus son patógenos emergentes y siguen representando una amenaza para la salud mundial, y tomando en cuenta que la vacunación es hasta ahora la mejor herramienta que tenemos para detener esta crisis sanitaria, es imperativo que se siga desarrollando la próxima generación de vacunas con el objetivo de disponer de mejores herramientas para enfrentar al virus y las variantes que vayan surgiendo. El trabajo aquí presentado aporta información importante que puede utilizarse para el desarrollo de nuevas vacunas.

## CONCLUSIÓN

Se identificaron epítomos de células T de las proteínas de SARS-CoV-2 que tienen un porcentaje de identidad mayor al 50% entre éste y los otros coronavirus humanos. Los porcentajes de identidad para estos epítomos de células T fueron considerablemente más altos que los encontrados para los epítomos de células B. Adicionalmente, la evidencia experimental generada por otros grupos de investigación sugiere que se generan respuestas importantes de reacción cruzada de anticuerpos y células T a epítomos comunes de coronavirus en respuesta a la infección por SARS-CoV-2 y sus variantes, y que las células T con reactividad cruzada circulan en la población previo a la exposición al SARS-CoV-2, por lo que la contribución de estas respuestas a la protección queda por investigar más a fondo para confirmar nuestra hipótesis. Esperamos que esta tesis contribuya a comprender las respuestas inmunitarias provocadas en respuesta a la infección y resalte el papel potencial de los epítomos de células T en la inmunidad y su influencia en el comportamiento de los brotes.

## **PERSPECTIVAS**

A partir de los estudios y discusión presentada en esta tesis, las perspectivas de trabajos futuros se orientan, en un primer término, al análisis y mapeo de las mutaciones del SARS-CoV-2 y las variantes de importancia clínica que se han descrito a lo largo de esta pandemia. Estos resultados y la evidencia experimental estimulan a seguir el estudio no sólo de la proteína S, sino de otras proteínas que conforman a este virus, además del estudio de sus regiones más conservadas con otros coronavirus humanos para la mejora de las vacunas actuales.

En un segundo término, se abordará de manera experimental el potencial de las respuestas celulares en el control de la infección por SARS-CoV-2 inducidos por los epítomos aquí predichos, que podrían ser tan importantes como la respuesta de anticuerpos, ganando así, relevancia en el diseño de vacunas contra el SARS-CoV-2 de próxima generación.

## BIBLIOGRAFÍA

1. Munster VJ, Koopmans M, van Doremalen N, van Riel D, de Wit E. A novel Coronavirus emerging in China - key questions for impact assessment. *N Engl J Med.* 2020;382(8):692–4.
2. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis.* 2020;20(5):533–4.
3. Singhal T. A review of Coronavirus disease-2019 (COVID-19). *Indian J Pediatr.* 2020;87(4):281–6.
4. Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, et al. Bats are natural reservoirs of SARS-like coronaviruses. *Science.* 2005;310(5748):676–9.
5. COVID-19 Vaccine Visualization [Internet]. Covid-19vaccinetracker.org. [cited 2021 Jun 1]. Available from: <https://www.covid-19vaccinetracker.org/>
6. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Velesler D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell.* 2020;181(2):281-292.e6.
7. Pavon. Inmunología molecular, celular y traslacional, 2ed [Internet]. Cld.bz. [cited 2021 Jun 1]. Available from: <https://user-biackli.cld.bz/Pavon-Inmunologia-molecular-celular-y-traslacional-2ed>
8. Edridge AWD, Kaczorowska J, Hoste ACR, Bakker M, Klein M, Jebbink MF, et al. Coronavirus protective immunity is short-lasting [Internet]. bioRxiv. 2020. Available from: <http://dx.doi.org/10.1101/2020.05.11.20086439>
9. Callow KA, Parry HF, Sergeant M, Tyrrell DAJ. The time course of the immune response to experimental coronavirus infection of man. *Epidemiol Infect.* 1990;105(2):435–46.
10. Guo X, Guo Z, Duan C, Chen Z, Wang G, Lu Y, et al. Long-term persistence of IgG antibodies in SARS-CoV infected healthcare workers [Internet]. bioRxiv. 2020. Available from: <http://dx.doi.org/10.1101/2020.02.12.20021386>
11. Cui J, Li F, Shi Z-L. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol.* 2019;17(3):181–92.

12. Yi, Y., Lagniton, P. N. P., Ye, S., Li, E., & Xu, R.-H. (2020). COVID-19: what has been learned and to be learned about the novel coronavirus disease. *International Journal of Biological Sciences*, 16(10), 1753–1766.
13. Mubarak, A., Alturaiki, W., & Hemida, M. G. (2019). Middle East respiratory syndrome Coronavirus (MERS-CoV): Infection, immunological response, and vaccine development. *Journal of Immunology Research*, 2019, 1–11.
14. Prasad A, Prasad M. SARS-CoV-2: the emergence of a viral pathogen causing havoc on human existence. *J Genet* [Internet]. 2020;99(1). Available from: <http://dx.doi.org/10.1007/s12041-020-01205-x>
15. *Taller de Actualización Bioquímica, Facultad de Medicina, UNAM*. (2019, October 29). Unam.mx. <http://bq.facmed.unam.mx/tab/>
16. Dhama K, Khan S, Tiwari R, Sircar S, Bhat S, Malik YS, et al. Coronavirus disease 2019-COVID-19. *Clin Microbiol Rev* [Internet]. 2020;33(4). Available from: <http://dx.doi.org/10.1128/CMR.00028-20>
17. Tang X, Wu C, Li X, Song Y, Yao X, Wu X, et al. On the origin and continuing evolution of SARS-CoV-2. *Natl Sci Rev*. 2020;7(6):1012–23.
18. Oliva Marín JE. SARS-CoV-2: origen, estructura, replicación y patogénesis. *Alerta* [Internet]. 2020;3(2). Available from: <http://dx.doi.org/10.5377/alerta.v3i2.9619>
19. Chen Y, Liu Q, Guo D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J Med Virol*. 2020;92(4):418–23.
20. Fehr AR, Perlman S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol*. 2015;1282:1–23.
21. Lam Cabanillas, E. R., León Risco, A. O., León Risco, K. B., Llamas Hoyos, G. L., López Zavaleta, R. M., Luzuriaga Tirado, E. del R., Mendoza Blas, A. M., & Huamán Saavedra, J. J. (2021). Molecular basis of COVID-19 pathogenesis and in silico studies of potential pharmacological treatment. *Revista de La Facultad de Medicina Humana*, 21(2), 417–432.
22. Casalino, L., Gaieb, Z., Goldsmith, J. A., Hjorth, C. K., Dommer, A. C., Harbison, A. M., Fogarty, C. A., Barros, E. P., Taylor, B. C., McLellan, J. S.,



- Fadda, E., & Amaro, R. E. (2020). Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein. *ACS Central Science*, 6(10), 1722–1734.
23. Ruiz-Bravo A, Jiménez-Valera M. SARS-CoV-2 y pandemia de síndrome respiratorio agudo (COVID-19). *Ars Pharm.* 2020;61(2):63–79.
24. *Modeling of the SARS-COV-2 Genome using D-I-TASSER.* (n.d.). Zhanggroup.Org. Retrieved November 13, 2021, from <https://zhanggroup.org//COVID-19/>.
25. Arya, R., Kumari, S., Pandey, B., Mistry, H., Bihani, S. C., Das, A., Prashar, V., Gupta, G. D., Panicker, L., & Kumar, M. (2021). Structural insights into SARS-CoV-2 proteins. *Journal of Molecular Biology*, 433(2), 166725.
26. Salimi-Jeda, A., Abbassi, S., Mousavizadeh, A., Esghaie, M., Bokharaei-Salim, F., Jeddi, F., Shafaati, M., & Abdoli, A. (2021). SARS-CoV-2: Current trends in emerging variants, pathogenesis, immune responses, potential therapeutic, and vaccine development strategies. *International Immunopharmacology*, 101(Pt A), 108232.
27. Debnath, P., Khan, U., & Khan, M. S. (2021). Characterization and structural prediction of ORF10, ORF7b, ORF7a, ORF6, membrane glycoprotein, and envelope protein in SARS-CoV-2 Bangladeshi variant through bioinformatics approach. En *bioRxiv*. <https://doi.org/10.1101/2021.10.01.452232>
28. Kushwaha SK, Kesarwani V, Choudhury S, Gandhi S, Sharma S. SARS-CoV-2 transcriptome analysis and molecular cataloguing of immunodominant epitopes for multi-epitope based vaccine design [Internet]. *bioRxiv*. 2020. Available from: <http://dx.doi.org/10.1101/2020.05.14.097170>
29. Chakraborty, C., Bhattacharya, M., & Sharma, A. R. (2022). Present variants of concern and variants of interest of severe acute respiratory syndrome coronavirus 2: Their significant mutations in S-glycoprotein, infectivity, re-infectivity, immune escape and vaccines activity. *Reviews in Medical Virology*, 32(2). <https://doi.org/10.1002/rmv.2270>
30. Paola TP Dra. Visión panorámica del sistema inmune. *Rev médica Clín Las Condes*. 2012;23(4):446–57.

31. Chabalgoity J, Pereira M, Rial A. Inmunidad contra los agentes infecciosos. *Temas Bacteriol y Virol Médica*. 2004;99–114.
32. Prompetchara E, Ketloy C, Palaga T. Immune responses in COVID-19 and potential vaccines: Lessons learned from SARS and MERS epidemic. *Asian Pacific Journal of Allergy and Immunology*. 2020;38(1):1-9. doi:10.12932/AP-200220-0772
33. Jeyanathan M, Afkhami S, Smaill F, Miller MS, Lichty BD, Xing Z. Immunological considerations for COVID-19 vaccine strategies. *Nature Reviews Immunology*. 2020;20(10):615-632. doi:10.1038/s41577-020-00434-6
34. Du Y, Tu L, Zhu P, et al. Clinical features of 85 fatal cases of COVID-19 from Wuhan: A retrospective observational study. *American Journal of Respiratory and Critical Care Medicine*. 2020;201(11):1372-1379. doi:10.1164/rccm.202003-0543OC
35. Kar T, Narsaria U, Basak S, Deb D, Castiglione F, Mueller DM, et al. A candidate multi-epitope vaccine against SARS-CoV-2. *Sci Rep*. 2020;10(1):10895.
36. Vernot, J. P. (1990). Mecanismos de procesamiento y presentación de antígeno en el sistema inmune. *Biomedica: Revista Del Instituto Nacional de Salud*, 10(1–4), 54.
37. KUBY. Inmunología, 7e [Internet]. Mhmedical.com. [cited 2021 Jun 1]. Available from: <https://accessmedicina.mhmedical.com/content.aspx?bookid=1953&sectionid=143392292>
38. den Haan JMM, Arens R, van Zelm MC. The activation of the adaptive immune system: cross-talk between antigen-presenting cells, T cells and B cells. *Immunol Lett*. 2014;162(2 Pt B):103–12.
39. Sohail, M. S., Ahmed, S. F., Quadeer, A. A., & McKay, M. R. (2021). In silico T cell epitope identification for SARS-CoV-2: Progress and perspectives. *Advanced Drug Delivery Reviews*, 171, 29–47. <https://doi.org/10.1016/j.addr.2021.01.007>

40. Montes-Grajales, D., & Olivero-Verbel, J. (2021). Bioinformatics prediction of SARS-CoV-2 epitopes as vaccine candidates for the Colombian population. *Vaccines*, 9(7), 797. <https://doi.org/10.3390/vaccines9070797>
41. Chen, H.-Z., Tang, L.-L., Yu, X.-L., Zhou, J., Chang, Y.-F., & Wu, X. (2020). Bioinformatics analysis of epitope-based vaccine design against the novel SARS-CoV-2. *Infectious Diseases of Poverty*, 9(1), 88. <https://doi.org/10.1186/s40249-020-00713-3>
42. Yang X, Yu X. An introduction to epitope prediction methods and software. *Rev Med Virol*. 2009;19(2):77–96.
43. Saha, S., & Raghava, G. P. S. (2004). BcePred: Prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. En *Lecture Notes in Computer Science* (pp. 197–204). Springer Berlin Heidelberg
44. Chin, C. F., Lai, J. Y., Choong, Y. S., Anthony, A. A., Ismail, A., & Lim, T. S. (2017). Delineation of B-cell Epitopes of Salmonella enterica serovar Typhi Hemolysin E: Potential antibody therapeutic target. *Scientific Reports*, 7(1), 2176. <https://doi.org/10.1038/s41598-017-01987-8>
45. Rammensee, H., Bachmann, J., Emmerich, N. P., Bachor, O. A., & Stevanović, S. (1999). SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*, 50(3–4), 213–219. <https://doi.org/10.1007/s002510050595>
46. Fleri W, Paul S, Dhanda SK, Mahajan S, Xu X, Peters B, et al. The Immune Epitope Database and Analysis Resource in Epitope discovery and synthetic vaccine design. *Front Immunol*. 2017;8:278.
47. Dhanda SK, Mahajan S, Paul S, Yan Z, Kim H, Jespersen MC, et al. IEDB-AR: immune epitope database-analysis resource in 2019. *Nucleic Acids Res*. 2019;47(W1):W502–6.
48. Jespersen, M. C., Peters, B., Nielsen, M., & Marcatili, P. (2017). BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Research*, 45(W1), W24–W29. <https://doi.org/10.1093/nar/gkx346>

49. Larsen JEP, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res.* 2006;2(1):2.
50. Hoof, I., Peters, B., Sidney, J., Pedersen, L. E., Sette, A., Lund, O., Buus, S., & Nielsen, M. (2009). NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics*, 61(1), 1–13. <https://doi.org/10.1007/s00251-008-0341-z>
51. Karosiene, E., Lundegaard, C., Lund, O., & Nielsen, M. (2012). NetMHCcons: a consensus method for the major histocompatibility complex class I predictions. *Immunogenetics*, 64(3), 177–186. <https://doi.org/10.1007/s00251-011-0579-8>
52. Nielsen, M., Lundegaard, C., Worning, P., Lauemøller, S. L., Lamberth, K., Buus, S., Brunak, S., & Lund, O. (2003). Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Science: A Publication of the Protein Society*, 12(5), 1007–1017. <https://doi.org/10.1110/ps.0239403>
53. Peters, B., & Sette, A. (2005). Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics*, 6(1), 132. <https://doi.org/10.1186/1471-2105-6-132>
54. Sidney, J., Assarsson, E., Moore, C., Ngo, S., Pinilla, C., Sette, A., & Peters, B. (2021, diciembre 26). Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. *NeuEve*. <https://www.neueve.com/blogs/immunome-research/4-1-2>
55. Chelly Dagdia, Z., Avdeyev, P., & Bayzid, M. S. (2021). Biological computation and computational biology: survey, challenges, and discussion. *Artificial Intelligence Review*, 54(6), 4169–4235.
56. Sievers, F., & Higgins, D. G. (2014). Clustal omega. *Et al [Current Protocols in Bioinformatics]*, 48(1), 3.13.1-16
57. *EMBOSS: needle manual*. (n.d.). Bioinformatics.NL. Retrieved November 13, 2021, from <https://www.bioinformatics.nl/cgi-bin/emboss/help/needle>

58. Greenbaum, J., Sidney, J., Chung, J., Brander, C., Peters, B., & Sette, A. (2011). Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. *Immunogenetics*, 63(6), 325–335.
59. Weiskopf, D., Angelo, M. A., de Azeredo, E. L., Sidney, J., Greenbaum, J. A., Fernando, A. N., Broadwater, A., Kolla, R. V., De Silva, A. D., de Silva, A. M., Mattia, K. A., Doranz, B. J., Grey, H. M., Shresta, S., Peters, B., & Sette, A. (2013). Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. *Proceedings of the National Academy of Sciences of the United States of America*, 110(22), E2046-53.
60. *The Allele Frequency Net Database - Allele, haplotype and genotype frequencies in Worldwide Populations.* (s/f). Allelefreqencies.net. Recuperado el 21 de noviembre de 2021, de <http://www.allelefreqencies.net/>
61. Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B, Sette A. A sequence homology and bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. *Cell Host Microbe*. 2020;27(4):671-680.e2.
62. Paul, S., Sidney, J., Sette, A., & Peters, B. (2016). TepiTool: A pipeline for computational prediction of T cell Epitope candidates. *Et al [Current Protocols in Immunology]*, 114(1), 18.19.1-18.19.24.
63. Woo, H., Park, S.-J., Choi, Y. K., Park, T., Tanveer, M., Cao, Y., Kern, N. R., Lee, J., Yeom, M. S., Croll, T. I., Seok, C., & Im, W. (2020). Developing a fully glycosylated full-length SARS-CoV-2 spike protein model in a viral membrane. *The Journal of Physical Chemistry. B*, 124(33), 7128–7137. <https://doi.org/10.1021/acs.jpcc.0c04553>
64. Denis M, Vandeweerd V, Verbeeke R, Van Der Vliet D. COVID-19 living paper: overview of information available to support the development of medical countermeasures and interventions against COVID-19. 2020; Available from: <http://dx.doi.org/10.5281/ZENODO.3765227>

65. Kumar S. Drug and vaccine design against novel Coronavirus (2019-nCoV) spike protein through computational approach [Internet]. Preprints. 2020. Available from: <http://dx.doi.org/10.20944/preprints202002.0071.v1>
66. Slathia P, Sharma P. Prediction of T and B cell epitopes in the proteome of SARS-CoV-2 for potential use in diagnostics and vaccine design [Internet]. ChemRxiv. 2020. Available from: <http://dx.doi.org/10.26434/chemrxiv.12116943.v1>
67. Ahmed SF, Quadeer AA, McKay MR. COVIDep: a web-based platform for real-time reporting of vaccine target recommendations for SARS-CoV-2. *Nat Protoc* (2020). doi: 10.1038/s41596-020-0358-9.
68. Qamar, M. T. ul, Shahid, F., Ashfaq, U. A., Aslam, S., Fatima, I., Fareed, M. M., Zohaib, A., & Chen, L.-L. (2020). Structural modeling and conserved epitopes prediction against SARS-COV-2 structural proteins for vaccine development. In *Research Square*. <https://doi.org/10.21203/rs.2.23973/v1>
69. Kiyotani K, Toyoshima Y, Nemoto K, Nakamura Y. Bioinformatic prediction of potential T cell epitopes for SARS-Cov-2. *J Hum Genet*. 2020;65(7):569–75.
70. Fast, E., Altman, R. B., & Chen, B. (2020). Potential T-cell and B-cell Epitopes of 2019-nCoV. In *bioRxiv*. <https://doi.org/10.1101/2020.02.19.955484>
71. Li, Y., Lai, D.-Y., Zhang, H.-N., Jiang, H.-W., Tian, X., Ma, M.-L., Qi, H., Meng, Q.-F., Guo, S.-J., Wu, Y., Wang, W., Yang, X., Shi, D.-W., Dai, J.-B., Ying, T., Zhou, J., & Tao, S.-C. (2020). Linear epitopes of SARS-CoV-2 spike protein elicit neutralizing antibodies in COVID-19 patients. *Cellular & Molecular Immunology*, 17(10), 1095–1097.
72. Poh, C. M., Carissimo, G., Wang, B., Amrun, S. N., Lee, C. Y.-P., Chee, R. S.-L., Fong, S.-W., Yeo, N. K.-W., Lee, W.-H., Torres-Ruesta, A., Leo, Y.-S., Chen, M. I.-C., Tan, S.-Y., Chai, L. Y. A., Kalimuddin, S., Kheng, S. S. G., Thien, S.-Y., Young, B. E., Lye, D. C., ... Ng, L. F. P. (2020). Two linear epitopes on the SARS-CoV-2 spike protein that elicit neutralising antibodies in COVID-19 patients. *Nature Communications*, 11(1), 2806.
73. Zhang, B.-Z., Hu, Y.-F., Chen, L.-L., Yau, T., Tong, Y.-G., Hu, J.-C., Cai, J.-P., Chan, K.-H., Dou, Y., Deng, J., Wang, X.-L., Hung, I. F.-N., To, K. K.-W.,

- Yuen, K. Y., & Huang, J.-D. (2020). Mining of epitopes on spike protein of SARS-CoV-2 from COVID-19 patients. *Cell Research*, *30*(8), 702–704.
74. Yuan, M., Wu, N. C., Zhu, X., Lee, C.-C. D., So, R. T. Y., Lv, H., Mok, C. K. P., & Wilson, I. A. (2020). A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science (New York, N.Y.)*, *368*(6491), 630–633.
75. Polyiam, K., Phoolcharoen, W., Butkhot, N., Srisaowakarn, C., Thitithanyanont, A., Auewarakul, P., Hoonsuwan, T., Ruengjitchatchawalya, M., Mekvichitsaeng, P., & Roshorm, Y. M. (2021). Immunodominant linear B cell epitopes in the spike and membrane proteins of SARS-CoV-2 identified by immunoinformatics prediction and immunoassay. *Scientific Reports*, *11*(1), 20383.
76. Amrun, S. N., Lee, C. Y.-P., Lee, B., Fong, S.-W., Young, B. E., Chee, R. S.-L., Yeo, N. K.-W., Torres-Ruesta, A., Carissimo, G., Poh, C. M., Chang, Z. W., Tay, M. Z., Chan, Y.-H., Chen, M. I.-C., Low, J. G.-H., Tambyah, P. A., Kalimuddin, S., Pada, S., Tan, S.-Y., ... Ng, L. F. P. (2020). Linear B-cell epitopes in the spike and nucleocapsid proteins as markers of SARS-CoV-2 exposure and disease severity. *EBioMedicine*, *58*(102911), 102911.
77. Le Bert, N., Tan, A. T., Kunasegaran, K., Tham, C. Y. L., Hafezi, M., Chia, A., Chng, M. H. Y., Lin, M., Tan, N., Linster, M., Chia, W. N., Chen, M. I.-C., Wang, L.-F., Ooi, E. E., Kalimuddin, S., Tambyah, P. A., Low, J. G.-H., Tan, Y.-J., & Bertoletti, A. (2020). SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature*, *584*(7821), 457–462.
78. Peng, Y., Mentzer, A. J., Liu, G., Yao, X., Yin, Z., Dong, D., Dejnirattisai, W., Rostron, T., Supasa, P., Liu, C., Lopez-Camacho, C., Slon-Campos, J., Zhao, Y., Stuart, D., Paeson, G., Grimes, J., Antson, F., Bayfield, O. W., Hawkins, D. E., ... Dong, T. (2020). Broad and strong memory CD4 + and CD8 + T cells induced by SARS-CoV-2 in UK convalescent COVID-19 patients. *BioRxiv.Org: The Preprint Server for Biology*. <https://doi.org/10.1101/2020.06.05.134551>

79. Kar, T., Narsaria, U., Basak, S., Deb, D., Castiglione, F., Mueller, D. M., & Srivastava, A. P. (2020). A candidate multi-epitope vaccine against SARS-CoV-2. *Scientific Reports*, *10*(1), 10895.
80. Shomuradova, A. S., Vagida, M. S., Sheetikov, S. A., Zornikova, K. V., Kiryukhin, D., Titov, A., Peshkova, I. O., Khmelevskaya, A., Dianov, D. V., Malasheva, M., Shmelev, A., Serdyuk, Y., Bagaev, D. V., Pivnyuk, A., Shcherbinin, D. S., Maleeva, A. V., Shakirova, N. T., Pilunov, A., Malko, D. B., ... Efimov, G. A. (2020). SARS-CoV-2 epitopes are recognized by a public and diverse repertoire of human T cell receptors. *Immunity*, *53*(6), 1245-1257.e5.
81. Chukwudozie, O. S., Gray, C. M., Fagbayi, T. A., Chukwuanukwu, R. C., Oyebanji, V. O., Bankole, T. T., Adewole, R. A., & Daniel, E. M. (2021). Immuno-informatics design of a multimeric epitope peptide based vaccine targeting SARS-CoV-2 spike glycoprotein. *PloS One*, *16*(3), e0248061.
82. ASANLAATK epitope - Immune Epitope Database (IEDB). (s/f). [iedb.org](http://www.iedb.org). Recuperado el 22 de noviembre de 2021, de <http://www.iedb.org/epitope/4321>
83. Hisham, Y., Ashhab, Y., Hwang, S.-H., & Kim, D.-E. (2021). Identification of highly conserved SARS-CoV-2 antigenic epitopes with wide coverage using reverse vaccinology approach. *Viruses*, *13*(5), 787.
84. D'Annessa, I., Marchetti, F., & Colombo, G. (2020). Differential antibody recognition by novel SARS-CoV-2 and SARS-CoV Spike protein receptor binding domains: Mechanistic insights and implications for the design of diagnostics and therapeutics. En *bioRxiv*. <https://doi.org/10.1101/2020.03.13.990267>
85. Kumar, S., Maurya, V. K., Prasad, A. K., Bhatt, M. L. B., & Saxena, S. K. (2020). Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV). *Virusdisease*, *31*(1), 13–21.
86. Stamatakis, G., Samiotaki, M., Mpakali, A., Panayotou, G., & Stratikos, E. (2020). Generation of SARS-CoV-2 S1 spike glycoprotein putative antigenic



- epitopes in vitro by intracellular aminopeptidases. *Journal of Proteome Research*, 19(11), 4398–4406.
87. Schreibing, F., Hannani, M., Ticconi, F., Fewings, E., Nagai, J. S., Begemann, M., Kuppe, C., Kurth, I., Kranz, J., Frank, D., Anslinger, T. M., Ziegler, P., Kraus, T., Enczmann, J., Balz, V., Windhofer, F., Balfanz, P., Kurts, C., Marx, G., ... Kramann, R. (2021). Dissecting CD8+ T cell pathology of severe SARS-CoV-2 infection by single-cell epitope mapping. En *bioRxiv* (p. 2021.03.03.432690). <https://doi.org/10.1101/2021.03.03.432690>
88. Pérez-Abeledo, M., & Sanz Moreno, J. C. (2021). Variantes de SARS-CoV-2, una historia todavía inacabada. *Vacunas*, 22(3), 173–179.
89. Wu, F., Wang, A., Liu, M., Wang, Q., Chen, J., Xia, S., Ling, Y., Zhang, Y., Xun, J., Lu, L., Jiang, S., Lu, H., Wen, Y., & Huang, J. (2020). Neutralizing antibody responses to SARS-CoV-2 in a COVID-19 recovered patient cohort and their implications. En *bioRxiv*. <https://doi.org/10.1101/2020.03.30.20047365>
90. Li, L., Zhao, Z., Yang, X., Li, W., Chen, S., Sun, T., Wang, L., He, Y., Liu, G., Han, X., Wen, H., Liu, Y., Chen, Y., Wang, H., Li, J., Su, Z., Du, C., Wang, Y., Li, X., ... Zhang, J. (2020). Identification of four linear B-cell epitopes on the SARS-CoV-2 spike protein able to elicit neutralizing antibodies. En *bioRxiv* (p.2020.12.13.422550). <https://doi.org/10.1101/2020.12.13.422550>
91. Lu, S., Xie, X.-X., Zhao, L., Wang, B., Zhu, J., Yang, T.-R., Yang, G.-W., Ji, M., Lv, C.-P., Xue, J., Dai, E.-H., Fu, X.-M., Liu, D.-Q., Zhang, L., Hou, S.-J., Yu, X.-L., Wang, Y.-L., Gao, H.-X., Shi, X.-H., ... Liu, R.-T. (2021). The immunodominant and neutralization linear epitopes for SARS-CoV-2. *Cell Reports*, 34(4), 108666.
92. Mishra, N., Huang, X., Joshi, S., Guo, C., Ng, J., Thakkar, R., Wu, Y., Dong, X., Li, Q., Pinapati, R. S., Sullivan, E., Caciula, A., Tokarz, R., Briese, T., Lu, J., & Lipkin, W. I. (2021). Immunoreactive peptide maps of SARS-CoV-2. *Communications Biology*, 4(1), 225.

93. Yi, Z., Ling, Y., Zhang, X., Chen, J., Hu, K., Wang, Y., Song, W., Ying, T., Zhang, R., Lu, H., & Yuan, Z. (2020). Functional mapping of B-cell linear epitopes of SARS-CoV-2 in COVID-19 convalescent population. *Emerging Microbes & Infections*, 9(1), 1988–1996.
94. Cruz, D. J. M., Kim, C.-J., & Shin, H.-J. (2006). Phage-displayed peptides having antigenic similarities with porcine epidemic diarrhea virus (PEDV) neutralizing epitopes. *Virology*, 354(1), 28–34.
95. Fan, B. E., Chong, V. C. L., Chan, S. S. W., Lim, G. H., Lim, K. G. E., Tan, G. B., Mucheli, S. S., Kuperan, P., & Ong, K. H. (2020). Hematologic parameters in patients with COVID-19 infection. *American Journal of Hematology*, 95(6), E131–E134.
96. Chen, Z., Ruan, P., Wang, L., Nie, X., Ma, X., & Tan, Y. (2021). T and B cell Epitope analysis of SARS-CoV-2 S protein based on immunoinformatics and experimental research. *Journal of Cellular and Molecular Medicine*, 25(2), 1274–1289. <https://doi.org/10.1111/jcmm.16200>
97. Bertoletti, A., Le Bert, N., Qui, M., & Tan, A. T. (2021). SARS-CoV-2-specific T cells in infection and vaccination. *Cellular & Molecular Immunology*, 18(10), 2307–2312.
98. Sekine, T., Perez-Potti, A., Rivera-Ballesteros, O., Strålin, K., Gorin, J.-B., Olsson, A., Llewellyn-Lacey, S., Kamal, H., Bogdanovic, G., Muschiol, S., Wullimann, D. J., Kammann, T., Emgård, J., Parrot, T., Folkesson, E., Rooyackers, O., Eriksson, L. I., Henter, J.-I., Sönnernborg, A., ... Unge, C. (2020). Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. *Cell*, 183(1), 158-168.e14.
99. Tan, C.-W., Chia, W.-N., Young, B. E., Zhu, F., Lim, B.-L., Sia, W.-R., Thein, T.-L., Chen, M. I.-C., Leo, Y.-S., Lye, D. C., & Wang, L.-F. (2021). Pan-sarbecovirus neutralizing antibodies in BNT162b2-immunized SARS-CoV-1 survivors. *The New England Journal of Medicine*, 385(15), 1401–1406. <https://doi.org/10.1056/NEJMoa2108453>
100. Kundu, R., Narean, J. S., Wang, L., Fenn, J., Pillay, T., Fernandez, N. D., Conibear, E., Koycheva, A., Davies, M., Tolosa-Wright, M., Hakki, S.,

- Varro, R., McDermott, E., Hammett, S., Cutajar, J., Thwaites, R. S., Parker, E., Rosadas, C., McClure, M., ... Lalvani, A. (2022). Cross-reactive memory T cells associate with protection against SARS-CoV-2 infection in COVID-19 contacts. *Nature Communications*, 13(1), 80. <https://doi.org/10.1038/s41467-021-27674-x>
101. Lipsitch, M., Grad, Y. H., Sette, A., & Crotty, S. (2020). Cross-reactive memory T cells and herd immunity to SARS-CoV-2. *Nature Reviews. Immunology*, 20(11), 709–713.
102. Sette, A., & Crotty, S. (2020). Pre-existing immunity to SARS-CoV-2: the knowns and unknowns. *Nature Reviews. Immunology*, 20(8), 457–458. <https://doi.org/10.1038/s41577-020-0389-z>
103. Sette, A., & Crotty, S. (2021). Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell*, 184(4), 861–880.
104. Guerrero, G., Picozza, M., D'Orso, S., Placido, R., Pirronello, M., Verdiani, A., Termine, A., Fabrizio, C., Giannessi, F., Sambucci, M., Balice, M. P., Caltagirone, C., Salvia, A., Rossini, A., Battistini, L., & Borsellino, G. (2021). BNT162b2 vaccination induces durable SARS-CoV-2 specific T cells with a stem cell memory phenotype. *Science Immunology*, eabl5344.
105. Heide, J., Schulte, S., Kohsar, M., Brehm, T. T., Herrmann, M., Karsten, H., Marget, M., Peine, S., Johansson, A. M., Sette, A., Lütgehetmann, M., Kwok, W. W., Sidney, J., & Schulze Zur Wiesch, J. (2021). Broadly directed SARS-CoV-2-specific CD4+ T cell response includes frequently detected peptide specificities within the membrane and nucleoprotein in patients with acute and resolved COVID-19. *PLoS Pathogens*, 17(9), e1009842. <https://doi.org/10.1371/journal.ppat.1009842>
106. Chakraborty, C., Bhattacharya, M., & Sharma, A. R. (2021). Present variants of concern and variants of interest of severe acute respiratory syndrome coronavirus 2: Their significant mutations in S-glycoprotein, infectivity, re-infectivity, immune escape and vaccines activity. *Reviews in Medical Virology*. <https://doi.org/10.1002/rmv.2270>

107. Faria, N. R., Mellan, T. A., Whittaker, C., Claro, I. M., Candido, D., Mishra, S., Crispim, M., Sales, F., Hawryluk, I., McCrone, J. T., Hulswit, R., Franco, L., Ramundo, M. S., de Jesus, J. G., Andrade, P. S., Coletti, T. M., Ferreira, G. M., Silva, C., Manuli, E. R., Pereira, R., ... Sabino, E. C. (2021). Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science (New York, N.Y.)*, eabh2644. Advance online publication. <https://doi.org/10.1126/science.abh2644>
108. Wang, P., Nair, M. S., Liu, L., Iketani, S., Luo, Y., Guo, Y., Wang, M., Yu, J., Zhang, B., Kwong, P. D., Graham, B. S., Mascola, J. R., Chang, J. Y., Yin, M. T., Sobieszczyk, M., Kyratsous, C. A., Shapiro, L., Sheng, Z., Huang, Y., & Ho, D. D. (2021). Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature*, 10.1038/s41586-021-03398-2. Advance online publication. <https://doi.org/10.1038/s41586-021-03398-2>
109. Zhou, D., Dejnirattisai, W., Supasa, P., Liu, C., Mentzer, A. J., Ginn, H. M., Zhao, Y., Duyvesteyn, H. M. E., Tuekprakhon, A., Nutalai, R., Wang, B., Paesen, G. C., Lopez-Camacho, C., Slon-Campos, J., Hallis, B., Coombes, N., Bewley, K., Charlton, S., Walter, T. S., ... Sreaton, G. R. (2021). Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell*, 184(9), 2348-2361.e6.
110. Lopez Bernal, J., Andrews, N., Gower, C., Gallagher, E., Simmons, R., Thelwall, S., Stowe, J., Tessier, E., Groves, N., Dabrera, G., Myers, R., Campbell, C. N. J., Amirthalingam, G., Edmunds, M., Zambon, M., Brown, K. E., Hopkins, S., Chand, M., & Ramsay, M. (2021). Effectiveness of Covid-19 vaccines against the B.1.617.2 (delta) variant. *The New England Journal of Medicine*, 385(7), 585–594.
111. Abdool Karim, S. S., & de Oliveira, T. (2021). New SARS-CoV-2 variants - clinical, public health, and vaccine implications. *The New England Journal of Medicine*, 384(19), 1866–1868.
112. Abu-Raddad, L. J., Chemaitelly, H., Butt, A. A., & National Study Group for COVID-19 Vaccination. (2021). Effectiveness of the BNT162b2 covid-19

vaccine against the B.1.1.7 and B.1.351 variants. *The New England Journal of Medicine*, 385(2), 187–189.

## ANEXOS

### Anexo 1. Nuevos epítomos lineales de células B y células T predichos en este estudio.

En el anexo 1 se presenta el compendio de epítomos nuevos predichos en este estudio de células B lineales (Tabla 7) y células T restringidos a MHC-I, MHC-II y promiscuos (Tabla 8) para cada una de las proteínas de SARS-CoV-2. Así como los epítomos reportados experimentalmente y capaces de montar una respuesta inmune (Tabla 9).

**Tabla 7.** Nuevos epítomos lineales de células B predichos en este estudio. Anexo 1.

Proteína	Secuencia	Posición	Fuente
S	DPFLGVYYHKNNKSWME	138-154	En este estudio
	MDLEGKQGNFKNL	177-189	En este estudio
	KHTPINLVRDLPQGFS	206-221	En este estudio
	NGTITD	282-287	En este estudio
	GDEVRQIAPGQTGKIADYNYKLPDD	404-428	En este estudio
	YQAGSTPCNGV	473-483	En este estudio
	YGFQPTNGVGYQ	495-506	En este estudio
	VITPGTNTSN	597-606	En este estudio
	NCTEVPVAIHADQLTPT	616-632	En este estudio
	RVYSTGNSVFAQ	634-644	En este estudio
	VNNSYECDIPI	656-666	En este estudio
	ASYQTQTNSPRRARSVASQ	672-690	En este estudio
	VEQDKNTQE	772-780	En este estudio
	KQIYKTPPIKDFGGF	786-800	En este estudio
	ILPDPSKPSKRS	805-816	En este estudio
	LADAGFIKQYGDCLG	828-842	En este estudio
RNFYEPQIITD	1107-1118	En este estudio	
M	YRIGNYKLNTDHSSSSDNIA	199-218	En este estudio
N	TGPEAGLPYGANK	115-127	En este estudio
	KLDDKDPNFKD	338-348	En este estudio
	SKQLQQSMSSADS	404-416	En este estudio
ORF1ab	DARTAPH	75-81	En este estudio
	QYGRSGE	96-102	En este estudio
	GNKGAGGHSYGA	127-138	En este estudio
	DELGTDPYEDFQENWNTKHSSG	147-168	En este estudio
	GGAYTR	179-184	En este estudio
	FCGPDGYPL	190-198	En este estudio
	ERSEKSYELQ	243-252	En este estudio
	PRVEKK	286-291	En este estudio
	PVASPNEC	305-312	En este estudio
	CGETSWQTG	326-334	En este estudio
	NLTKEGATTCGYL	348-360	En este estudio
	HNSEVGPEHSLAEYHNES	374-391	En este estudio
	GVVGESEGL	437-446	En este estudio
	VTKGKAKKGAWNIG	511-524	En este estudio
	LQPLEQPTSEAVEAPLV	763-779	En este estudio
	TLKGGAPTKVTFG	814-826	En este estudio
	YPPDEDEEEGDCEEEEFEPST	925-945	En este estudio

TSAALQPEEEQEEDWLDDDSQ	965-985	En este estudio
IVEEAKVK	1045-1053	En este estudio
GGVAGALNKATNNAMQVESDDYIATNGPLKV	1069-1099	En este estudio
PNVKNKE	1120-1126	En este estudio
EKQVEQ	1196-1201	En este estudio
AEIPKEEVKPFITESKPSVEQRKQDDKK	1204-1231	En este estudio
NLHPDSAT	1262-1269	En este estudio
TKKAGGTTE	1303-1311	En este estudio
KVPTDNYITTYPGQLNGYTVEE	1319-1341	En este estudio
TVSVSSPDAVTA	1474-1485	En este estudio
LTSSSKTPEEH	1490-1500	En este estudio
YKDWSYSGQSTQ	1511-1522	En este estudio
KSVYYTSNPTTFH	1533-1545	En este estudio
GQQFGPTYLDGADVTKIKPHNSHEG	1591-1615	En este estudio
YHTTDPSF	1635-1642	En este estudio
KWKYPQVN	1655-1662	En este estudio
NPPALQD	1691-1697	En este estudio
YRARAGE	1700-1706	En este estudio
GELGDV	1723-1728	En este estudio
GQQQTTL	1756-1762	En este estudio
APPAQYEL	1809-1816	En este estudio
EYTGNYQ	1826-1832	En este estudio
SSEYKGPI	1856-1863	En este estudio
NSYTTTIK	1871-1878	En este estudio
DPKLDNYKKDNSYFTE	1893-1909	En este estudio
IDLVPNQYPNASF	1912-1925	En este estudio
LTGYKKPASR	1944-1953	En este estudio
HYTPSFK	1974-1980	En este estudio
ATNKATYKP	1997-2005	En este estudio
KPVETSN	2017-2023	En este estudio
SEDAQGM	2030-2036	En este estudio
LKPVSEEVVENPTIQ	2044-2058	En este estudio
NSRIKAS	2187-2193	En este estudio
PTTIKAKN	2195-2201	En este estudio
SLDTYPSL	2297-2304	En este estudio
PINPTDQSSY	2480-2489	En este estudio
AGQKTYE	2508-2514	En este estudio
ANNTKGS	2529-2535	En este estudio
KSKCEESSAKSA	2546-2557	En este estudio
DVGDSAE	2579-2585	En este estudio
GFVDSDEVTK	2640-2649	En este estudio
DIEVTGDSC	2662-2670	En este estudio
VENMTPRD	2680-2687	En este estudio
IDGGVTRDIAS	2812-2822	En este estudio
SQRGGSYTNDK	2839-2849	En este estudio
DASGKVPYCY	2924-2933	En este estudio
TCERSE	2988-2993	En este estudio
AMDTTSYR	3220-3227	En este estudio
DFSNSGSDVLYQPQTSI	3240-3257	En este estudio
FPSGKV	3271-3276	En este estudio
DTANPKTPKY	3355-3364	En este estudio
NGSPSGV	3382-3388	En este estudio
GVHAGTDLEGN	3433-3443	En este estudio
VDRQTAQAAGTDTT	3449-3462	En este estudio

VYDDGAR	3700-3706	En este estudio
GLLPPKNSI	3827-3835	En este estudio
FATAQEAYEQAVANGDS	3957-3973	En este estudio
SEFDRD	3989-3994	En este estudio
KQARSEDKRAKV	4014-4025	En este estudio
YNTYKNTCDGTTF	4077-4089	En este estudio
SMDNSPNL	4115-4122	En este estudio
AAGTTQTACTIONDDN	4155-4167	En este estudio
FPKSDGTGTIYTELEP	4196-4211	En este estudio
RFVTDTPKGP	4214-4224	En este estudio
GNATEVPAN	4255-4263	En este estudio
LASGGQPIT	4284-4292	En este estudio
TGTGQAITVTPKANMDQESFG	4302-4322	En este estudio
DHPNPKGF	4335-4342	En este estudio
TTCANDPVG	4354-4362	En este estudio
LTPCGTGTSTD	4411-4421	En este estudio
QEKDEDDNL	4449-4457	En este estudio
IQTTPGSGVPV	4615-4626	En este estudio
SHVDTDLTK	4647-4655	En este estudio
FPPTSFG	4713-4719	En este estudio
VKPGNFNK	4802-4809	En este estudio
GNAAIS	4838-4843	En este estudio
MSYEDQ	4911-4916	En este estudio
SDVENP	4999-5004	En este estudio
YPKCDRA	5011-5017	En este estudio
YVKPGGTSSGDATTAY	5066-5081	En este estudio
ETDLTKGPHEF	5194-5204	En este estudio
DDYVYLPYPDPSRI	5216-5229	En este estudio
PLTKHPNQEY	5260-5269	En este estudio
DNTSRYWEP	5302-5310	En este estudio
CNAPGCDVT	5374-5382	En este estudio
KSHKPPI	5397-5403	En este estudio
CVGSDNVT	5421-5428	En este estudio
VGKPRPPLN	5493-5501	En este estudio
TFEKGDYGDA	5523-5532	En este estudio
STLQGGPGTGKS	5602-5613	En este estudio
GDPAQLPAP	5724-5732	En este estudio
KAHKDKS	5786-5792	En este estudio
SSAINR	5809-5814	En este estudio
YNSQNA	5839-5844	En este estudio
TQTVDSQSEY	5854-5865	En este estudio
HPTQAPTHLS	5944-5953	En este estudio
PGIPKD	5968-5973	En este estudio
CHATREAVG	6019-6027	En este estudio
TGYVDTPNNTDFSRVSAKPPPGDQF	6047-6071	En este estudio
IGNPKA	6257-6262	En este estudio
AQPCSDKAY	6278-6286	En este estudio
THSDKFT	6297-6303	En este estudio
LPGCDGGS	6336-6343	En este estudio
DSPCESHGKQ	6374-6383	En este estudio
HFDGQQGEV	6466-6474	En este estudio
KRDAPAH	6541-6547	En este estudio
IAKKPTET	6559-6566	En este estudio
VKGLQPSVGPQASL	6600-6614	En este estudio



	LQEFKPRSQ	6652-6660	En este estudio
	AQTGSS	6735-6740	En este estudio
	KLQSSQAWQPGVA	6796-6808	En este estudio
	QNYGDSATL	6826-6834	En este estudio
	AGSDKGVAPGT	6870-6880	En este estudio
	DFVSDADST	6900-6908	En este estudio
	YDPKTKNVTKENDSKE	6930-6945	En este estudio
	GKPREQI	7011-7017	En este estudio
ORF3a	QGEIKDATPSDFVR	17-30	En este estudio
	KIITLKKRWQL	61-71	En este estudio
	SKNPLL	135-140	En este estudio
	PYNSVT	159-164	En este estudio
	TSGDGTTSPISEHDYQIGGYTEKWESGV	170-197	En este estudio
	STQLSTDTGV	216-225	En este estudio
	VDEPEEHVQIHTIDGSSGVVNPVME	237-261	En este estudio
	EPIYDEPTTTTSVP	261-274	En este estudio
ORF6	LTENKYSQLDEEQP	44-57	En este estudio
ORF7a	LYHYQECVR	17-25	En este estudio
	EPCSSGTYEGNSPFHPLAD	33-51	En este estudio
	VKHVYQLRARSVSPKLFIRQEEVQEL	71-96	En este estudio
ORF7b	ELQDHNE	33-39	En este estudio
ORF8	QSCTQHQPYYVDDPCPIHFYSKW	23-45	En este estudio
	RVGARKSAP	48-56	En este estudio
	DEAGSKSPIQYIDIGN	63-78	En este estudio
	EDFLEY	106-111	En este estudio

**Tabla 8.** Nuevos epítomos de células T restringidos a MHC-I, MHC-II y promiscuos predichos en este estudio. Anexo 1.

Proteína	Secuencia	Posición	MHC	HLA	Fuente
S	RFDNPVLPF	78-86	I	HLA-A*23:01, HLA-A*24:02	En este estudio
	FQFCNDPFL	133-141	I	HLA-A*02:01, HLA-A*02:06	En este estudio
	LEPLVDLPI	223-231	I	HLA-B*40:01	En este estudio
	NVYADSFVI	394-402	I	HLA-A*02:01	En este estudio
	YQDVNCTEV	612-620	I	HLA-A*02:01	En este estudio
	TLADAGFIK	827-835	I	HLA-A*11:01, HLA-A*03:01	En este estudio
	QMAYRFNGI	901-909	I	HLA-B*08:01	En este estudio
	NQFNSAIGK	925-933	I	HLA-A*11:01	En este estudio
	AQALNTLVK	956-964	I	HLA-A*11:01	En este estudio
	AEVQIDRLI	989-997	I	HLA-B*40:01	En este estudio
	RAAEIRASANL	1014-1024	I	HLA-B*07:02, HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	En este estudio
	FPREGVFVS	1089-1097	I	HLA-B*35:01, HLA-B*07:02	En este estudio
	FVSNQTHWFVTQRNF	1095-1109	I	HLA-A*03:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01	En este estudio
	NASVVNIQK	1173-1181	I	HLA-A*11:01	En este estudio
M	SSDNIALLV	213-221	I	HLA-A*01:01	En este estudio
N	LPNNTASWF	45-53	I	HLA-B*35:01, HLA-B*07:02	En este estudio
	RIRGGDGKM	93-101	I	HLA-B*07:02	En este estudio
	KHWPQIAQF	299-307	I	HLA-A*23:01	En este estudio
	ASAFFGMSR	311-319	I	HLA-A*11:01, HLA-A*03:01	En este estudio
	GMSRIGMEV	316-324	I	HLA-A*02:01	En este estudio
	KTFPPTPEK	361-369	I	HLA-A*11:01, HLA-A*03:01	En este estudio
ORF1ab	SLVPGFNEK	3-11	I	HLA-A*03:01, HLA-A*11:01	En este estudio
	VLQVRDVLV	20-28	I	HLA-A*02:01	En este estudio
	VLSEARQHLK	38-47	I	HLA-A*02:01, HLA-A*11:01	En este estudio
	GLVEVEKGV	52-60	I	HLA-A*02:01	En este estudio
	RTAPHGHVM	77-85	I	HLA-B*07:02	En este estudio
	VMVELVAEL	84-92	I	HLA-A*02:01	En este estudio
	AELEGIQYG	90-98	I	HLA-B*44:02, HLA-B*44:03	En este estudio
	TLGVLVPHV	103-111	I	HLA-A*02:01	En este estudio
	VLVPHVGEIPVAY	106-118	I	HLA-A*02:01, HLA-B*07:02, HLA-B*35:01	En este estudio
	YPLECIKDL	196-204	I	HLA-B*07:02, HLA-B*35:01	En este estudio
	TLSEQLDFI	214-222	I	HLA-A*02:01	En este estudio
	REHEHEIAW	232-240	I	HLA-B*44:02, HLA-B*44:03	En este estudio
	YELQTPFEI	249-257	I	HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	En este estudio
	KTIQPRVEK	282-290	I	HLA-A*11:01, HLA-A*03:01	En este estudio
	KLDGFMGRIRSVY	292-304	I	HLA-A*02:01, HLA-B*35:01	En este estudio
	CEFCGTENL	341-349	I	HLA-B*40:01	En este estudio
	KEGATTCGY	351-359	I	HLA-B*44:02, HLA-B*44:03	En este estudio
	SEVGPEHSL	376-384	I	HLA-B*40:01	En este estudio
	TIAFGGCVF	403-411	I	HLA-B*35:01	En este estudio
FSYVGCHNK	411-419	I	HLA-A*11:01, HLA-A*03:01	En este estudio	

SEGLNDNLLEI	443-453	I	HLA-A*02:01, HLA-B*40:01	En este estudio
NINIVGDFK	460-468	I	HLA-A*11:01	En este estudio
KLNEEIAII	468-476	I	HLA-A*02:01	En este estudio
STSAFVETVK	483-492	I	HLA-A*02:01, HLA-A*11:01	En este estudio
GAWNIGEQQK	519-527	I	HLA-A*11:01	En este estudio
ILSPLYAFA	529-537	I	HLA-A*02:01	En este estudio
AARVRSIFSR	540-550	I	HLA-A*03:01, HLA-A*11:01, HLA-B*07:02	En este estudio
YSLRLIDAM	576-584	I	HLA-B*35:01	En este estudio
LATNNLVVM	590-598	I	HLA-B*35:01	En este estudio
VMAYITGGVVQL	597-608	I	HLA-A*02:01	En este estudio
WLTNIFGTVYEKL	612-624	I	HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01	En este estudio
KLKPVLDWL	623-631	I	HLA-A*02:01	En este estudio
FLRDGWEIV	641-649	I	HLA-A*02:01	En este estudio
CEIVGGQIV	657-665	I	HLA-B*40:01	En este estudio
ALCADSIII	687-695	I	HLA-A*02:01	En este estudio
ALNLGETFV	702-710	I	HLA-A*02:01	En este estudio
FLEGETLPTEVL	741-752	I	HLA-A*02:01, HLA-B*40:01	En este estudio
TEVLTEEVVLK	749-759	I	HLA-A*03:01, HLA-A*11:01, HLA-B*40:01	En este estudio
SEAVEAPLV	771-779	I	HLA-B*40:01	En este estudio
TPVCINGLM	781-789	I	HLA-B*07:02, HLA-B*35:01	En este estudio
KYCALAPNM	798-806	I	HLA-A*23:01	En este estudio
NMMVTNNTFTLK	805-816	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01	En este estudio
TLKGGAPTK	814-822	I	HLA-A*03:01, HLA-A*11:01	En este estudio
FGDDTVIEVQGYK	825-837	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01	En este estudio
YTVELGTEV	860-868	I	HLA-A*02:01	En este estudio
CVVADAVIK	873-881	I	HLA-A*11:01	En este estudio
SELLTPLGIDL	887-897	I	HLA-A*02:01, HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	En este estudio
YLFDESGEF	906-914	I	HLA-B*35:01, HLA-A*02:01	En este estudio
TTIQTIVEV	1000-1008	I	HLA-A*02:01	En este estudio
LEMELTPVV	1012-1020	I	HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	En este estudio
KLTDNVYIK	1033-1041	I	HLA-A*03:01, HLA-A*11:01	En este estudio
KPTVVVNAANVYLK	1053-1066	I	HLA-A*02:01, HLA-A*11:01, HLA-B*07:02, HLA-B*35:01	En este estudio
YIATNGPLK	1090-1098	I	HLA-A*11:01, HLA-A*03:01	En este estudio
VLSGHNLAK	1104-1112	I	HLA-A*03:01, HLA-A*11:01	En este estudio
CLHVVGPNVVK	1114-1124	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01	En este estudio
YENFNQHEV	1135-1143	I	HLA-B*40:01	En este estudio
HEVLLAPLL	1141-1149	I	HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	En este estudio
SLRVCVDTV	1161-1169	I	HLA-A*02:01	En este estudio
NVYLAVDFKNL	1172-1182	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01	En este estudio

KIKACVEEV	1231-1239	I	HLA-A*02:01	En este estudio
VTTTLEETK	1239-1247	I	HLA-A*11:01	En este estudio
KFLTENLLLY	1247-1256	I	HLA-A*01:01, HLA-A*23:01, HLA-B*35:01	En este estudio
LYIDINGNL	1255-1263	I	HLA-A*23:01	En este estudio
LVSDIDITFLK	1270-1280	I	HLA-A*11:01, HLA-B*35:01	En este estudio
APYIVGDVV	1283-1291	I	HLA-B*07:02	En este estudio
QEGVLTAVV	1292-1300	I	HLA-B*40:01	En este estudio
VLTAUVIPTKK	1295-1305	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01	En este estudio
PTDNYITTY	1321-1329	I	HLA-A*01:01	En este estudio
SAFYILPSI	1351-1359	I	HLA-A*02:01	En este estudio
QEILGTVSW	1365-1373	I	HLA-B*44:02, HLA-B*44:03	En este estudio
KLMPVCVETKA	1387-1397	I	HLA-A*02:01, HLA-B*35:01	En este estudio
KAIVSTIQR	1396-1404	I	HLA-A*11:01	En este estudio
VVDYGARFY	1415-1423	I	HLA-A*01:01	En este estudio
SLINTLNDLNETL	1433-1445	I	HLA-A*02:01	En este estudio
MPLGYVTHG	1448-1456	I	HLA-B*35:01	En este estudio
DAVTAYNGY	1481-1489	I	HLA-B*35:01	En este estudio
TQLGIEFLK	1521-1529	I	HLA-A*11:01	En este estudio
GEVITFDNLK	1548-1557	I	HLA-A*11:01, HLA-B*40:01	En este estudio
SLREVRTIKVFTTV	1561-1574	I	HLA-A*02:01, HLA-A*03:01, HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	En este estudio
YGQQFGPTY	1590-1598	I	HLA-B*35:01	En este estudio
LPNDDTLRV	1621-1629	I	HLA-B*35:01	En este estudio
YYHTTDPSPF	1634-1642	I	HLA-A*23:01, HLA-A*24:02	En este estudio
YMSALNHTK	1646-1654	I	HLA-A*11:01, HLA-A*03:01	En este estudio
YPQVNGLTS	1658-1666	I	HLA-B*35:01	En este estudio
YLATALLTL	1675-1683	I	HLA-A*02:01	En este estudio
ELKFNPPAL	1687-1695	I	HLA-B*08:01	En este estudio
ALQDAYYRA	1694-1702	I	HLA-A*02:01	En este estudio
ALILAYCNK	1712-1720	I	HLA-A*11:01, HLA-A*03:01	En este estudio
SYLFQHANL	1733-1741	I	HLA-A*23:01, HLA-A*24:02	En este estudio
CASEYTGNY	1823-1831	I	HLA-B*35:01	En este estudio
TLYCIDGALLTK	1844-1855	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02	En este estudio
VVCTEIDPK	1887-1895	I	HLA-A*11:01	En este estudio
YPNASFDNFKFV	1920-1931	I	HLA-A*02:01, HLA-B*07:02, HLA-B*35:01	En este estudio
FDDLNGDVV	1960-1968	I	HLA-B*35:01	En este estudio
TYKPNTWCI	2002-2010	I	HLA-A*23:01, HLA-A*24:02	En este estudio
CIRCLWSTK	2009-2017	I	HLA-A*03:01	En este estudio
VVENPTIQK	2051-2059	I	HLA-A*03:01, HLA-A*11:01	En este estudio
TEVVGDIILK	2069-2078	I	HLA-A*11:01, HLA-B*40:01	En este estudio
TEEVGHTDLM	2087-2096	I	HLA-B*40:01	En este estudio
VPWDTIANY	2133-2141	I	HLA-B*35:01	En este estudio
STTTNIVTR	2151-2159	I	HLA-A*11:01	En este estudio

NYMPYFFTLLL	2167-2177	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*07:02, HLA-B*08:01, HLA-B*35:01	En este estudio
ASMPPTIAK	2192-2200	I	HLA-A*03:01, HLA-A*11:01	En este estudio
FCLEASFNYL	2209-2218	I	HLA-A*02:01, HLA-B*35:01	En este estudio
LINIIWFLLSV	2226-2238	I	HLA-A*02:01	En este estudio
STNVTIATY	2273-2281	I	HLA-A*01:01, HLA-A*11:01, HLA-B*35:01	En este estudio
YPSLETIQI	2301-2309	I	HLA-B*35:01	En este estudio
SFKWDLTAF	2313-2321	I	HLA-A*23:01, HLA-A*24:02	En este estudio
TAFGLVAEWF	2319-2328	I	HLA-A*23:01, HLA-B*35:01	En este estudio
GLVAEWFLAYILF	2322-2334	I	HLA-A*01:01, HLA-A*02:01, HLA-A*11:01, HLA-A*03:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01, HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	En este estudio
AYILFTRFFYVL	2330-2341	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-B*08:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01	En este estudio
FYVLGLAAIM	2338-2347	I	HLA-A*23:01, HLA-A*24:02, HLA-B*35:01	En este estudio
IMQLFFSYFAVHFI	2346-2359	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02	En este estudio
WLMWLIINLVQM	2363-2374	I	HLA-A*02:01	En este estudio
AMVRMYIFFASFYVVK	2379-2395	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-B*08:01	En este estudio
YKRNRATRV	2413-2421	I	HLA-B*08:01	En este estudio
TTIVNGVRR	2424-2432	I	HLA-A*11:01	En este estudio
SYIVDSVTV	2488-2496	I	HLA-A*24:02, HLA-A*23:01	En este estudio
GSIHLYFDK	2499-2507	I	HLA-A*03:01, HLA-A*11:01	En este estudio
SLPINVIVF	2535-2543	I	HLA-A*23:01	En este estudio
KMFDAYVNTF	2589-2598	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01	En este estudio
FSSTFNVPM	2598-2606	I	HLA-B*35:01	En este estudio
ATAEELAK	2614-2622	I	HLA-A*03:01, HLA-A*11:01	En este estudio
VLSTFISAAR	2629-2638	I	HLA-A*02:01, HLA-A*11:01	En este estudio
QVVNVVTTK	2749-2757	I	HLA-A*03:01, HLA-A*11:01	En este estudio
KVTLVFLFV	2775-2783	I	HLA-A*02:01	En este estudio
TLVFLFVAAIFYL	2777-2789	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02	En este estudio
FVAAIFYLITPVHVM	2782-2796	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01	En este estudio
FSSEIIGYKAI	2802-2812	I	HLA-A*03:01, HLA-A*11:01, HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	En este estudio
STDTCFANK	2822-2830	I	HLA-A*03:01, HLA-A*11:01	En este estudio
REVGFFVPLPGTIL	2860-2874	I	HLA-A*02:01, HLA-B*40:01, HLA-B*07:02	En este estudio
FSAVGNICY	2889-2897	I	HLA-A*01:01, HLA-B*35:01	En este estudio
KLIEYTDFA	2901-2909	I	HLA-A*02:01	En este estudio
LAAECTIFK	2915-2923	I	HLA-A*03:01, HLA-A*11:01	En este estudio
NVLEGSVAY	2937-2945	I	HLA-B*35:01	En este estudio

SLRPDTRYVLM	2947-2957	I	HLA-A*02:01, HLA-B*07:02	En este estudio
SIIQFPNTY	2960-2968	I	HLA-B*35:01, HLA-A*11:01	En este estudio
TYLEGSVRVV	2967-2976	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02	En este estudio
VVTTFDSEY	2975-2983	I	HLA-B*35:01	En este estudio
SLPGVFCGVDAV	3013-3024	I	HLA-A*02:01	En este estudio
MFTPLIQPI	3030-3038	I	HLA-A*23:01, HLA-A*24:02	En este estudio
IVAGGIVAI	3047-3055	I	HLA-A*02:01	En este estudio
IVVTCLAYYFMRFR	3055-3068	I	HLA-A*11:01	En este estudio
YFMRFRRAFG	3063-3072	I	HLA-B*08:01, HLA-A*23:01, HLA-A*24:02	En este estudio
VVAFNTLLFL	3077-3086	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01	En este estudio
LMSFTVLCL	3086-3094	I	HLA-A*02:01	En este estudio
LPGVSVIY	3101-3109	I	HLA-B*35:01	En este estudio
VYSVIYLYLTFYLT	3104-3117	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02	En este estudio
VPFWITIAYIICI	3136-3148	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01	En este estudio
ISTKHFYWFFSNYL	3148-3161	I	HLA-A*23:01, HLA-A*24:02	En este estudio
FEEAALCTFLLNK	3175-3187	I	HLA-A*03:01, HLA-A*11:01, HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	En este estudio
TQYNRYLAL	3202-3210	I	HLA-B*08:01	En este estudio
MLNPNYEDL	3312-3320	I	HLA-A*02:01	En este estudio
QLRVIGHSM	3337-3345	I	HLA-B*07:02, HLA-B*08:01	En este estudio
AMRPNFTIK	3392-3400	I	HLA-A*03:01, HLA-A*11:01	En este estudio
GTDLEGIFY	3437-3445	I	HLA-A*01:01	En este estudio
FLNRFTTTLNDFNLVA	3482-3497	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*08:01	En este estudio
LLLTILTSL	3583-3592	I	HLA-A*02:01, HLA-B*08:01	En este estudio
STQWSLFFFL	3597-3606	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02	En este estudio
FFFLYENAFLPF	3603-3614	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01	En este estudio
FLPFAMGIIAMSA	3611-3623	I	HLA-A*02:01, HLA-B*35:01	En este estudio
FLCLFLLPSLATVAYFNM V	3635-3653	I	HLA-A*02:01, HLA-B*35:01	En este estudio
VYMPASWVM	3653-3661	I	HLA-A*23:01, HLA-A*24:02	En este estudio
WVMRIMTWLDMV	3659-3670	I	HLA-A*02:01	En este estudio
YASAVLLI	3685-3693	I	HLA-A*02:01, HLA-B*35:01	En este estudio
ILMTARTVYDDGARRVW TL	3693-3711	I	HLA-A*03:01, HLA-A*11:01, HLA-B*08:01, HLA-B*35:01	En este estudio
SMWALIISV	3732-3740	I	HLA-A*02:01	En este estudio
YSGVVTTVMFLARGIVF	3744-3760	I	HLA-A*02:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-B*08:01, HLA-B*35:01	En este estudio
FMCVEYCIFFITGNTL	3760-3776	I	HLA-A*02:01, HLA-A*23:01, HLA-B*40:01, HLA-B*44:03	En este estudio
MLVYCFLGY	3780-3788	I	HLA-A*03:01, HLA-B*35:01	En este estudio
YFCTCYFGL	3788-3796	I	HLA-A*23:01	En este estudio
CYFGLFCLLNRYFRLTLG V	3792-3810	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*08:01	En este estudio

LVSTQEFRY	3814-3822	I	HLA-A*01:01	En este estudio
QLHNDILLA	3893-3901	I	HLA-A*02:01	En este estudio
LAKDTTEAF	3900-3908	I	HLA-B*35:01	En este estudio
MQGAVDINK	3921-3929	I	HLA-A*11:01	En este estudio
LPSYAAFAT	3951-3959	I	HLA-B*35:01	En este estudio
SEFDRDAAM	3989-3997	I	HLA-B*40:01	En este estudio
ALNNIINNA	4044-4052	I	HLA-A*02:01	En este estudio
IIPLTAAKL	4061-4070	I	HLA-A*11:01, HLA-B*07:02	En este estudio
KLMVVIPDYNTYK	4069-4081	I	HLA-A*03:01, HLA-A*11:01, HLA-B*35:01	En este estudio
TCDGTTFTY	4083-4091	I	HLA-A*01:01, HLA-B*35:01	En este estudio
TYASALWEI	4090-4098	I	HLA-A*23:01, HLA-A*24:02	En este estudio
SPNLAWPLI	4119-4127	I	HLA-B*07:02	En este estudio
CTDDNALAYNTTK	4163-4176	I	HLA-A*01:01, HLA-A*03:01, HLA-A*11:01, HLA-B*35:01	En este estudio
KSDGTGTIY	4198-4206	I	HLA-A*01:01	En este estudio
TELEPPCRF	4207-4215	I	HLA-B*44:03, HLA-B*44:02	En este estudio
KVKYLYFIKGL	4224-4234	I	HLA-A*03:01, HLA-A*11:01, HLA-A*23:01	En este estudio
LQAGNATEV	4252-4260	I	HLA-A*02:01	En este estudio
CAFAVDAAK	4270-4278	I	HLA-A*11:01	En este estudio
GTSTDVVYR	4417-4425	I	HLA-A*11:01	En este estudio
VVYRAFDIYNDK	4422-4433	I	HLA-A*03:01, HLA-A*11:01	En este estudio
VAGFAFLK	4434-4442	I	HLA-A*11:01	En este estudio
NLIDSYFVVK	4456-4465	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01	En este estudio
HEETIYNLL	4474-4482	I	HLA-B*40:01	En este estudio
AVAKHDFFK	4487-4495	I	HLA-A*11:01, HLA-A*03:01	En este estudio
RVYANLGERV	4565-4574	I	HLA-A*03:01, HLA-A*11:01, HLA-A*24:02	En este estudio
NWYDFGDFI	4607-4615	I	HLA-A*23:01	En este estudio
SGVPVDSY	4621-4629	I	HLA-B*35:01	En este estudio
MPILTLTRA	4634-4642	I	HLA-B*35:01	En este estudio
AESHVDTDL	4645-4653	I	HLA-B*40:01	En este estudio
KPYIKWDLKLY	4655-4665	I	HLA-B*07:02, HLA-B*35:01	En este estudio
KLFDRYFKYWDQTY	4673-4686	I	HLA-A*03:01, HLA-A*11:01, HLA-B*35:01	En este estudio
VLFSTVFPP	4707-4715	I	HLA-A*02:01	En este estudio
FPPTSFGPLVRK	4713-4724	I	HLA-A*03:01, HLA-A*11:01, HLA-B*07:02, HLA-B*35:01	En este estudio
VPFVSTGYHFR	4730-4741	I	HLA-A*11:01, HLA-B*35:01	En este estudio
LLVYAADPAM	4763-4772	I	HLA-A*02:01, HLA-B*35:01	En este estudio
SVAALTNNVAF	4789-4799	I	HLA-A*02:01, HLA-B*35:01	En este estudio
TVKPGNFNK	4801-4809	I	HLA-A*11:01, HLA-A*03:01	En este estudio
FAQDGNAAISDYDY	4834-4847	I	HLA-B*35:01	En este estudio
YYRYNLPTM	4847-4855	I	HLA-A*23:01, HLA-A*24:02	En este estudio
QLLFVVEVDKY	4860-4871	I	HLA-A*11:01, HLA-B*35:01	En este estudio
QVIVNNLDK	4884-4892	I	HLA-A*11:01	En este estudio
KSAGFPFNK	4892-4900	I	HLA-A*03:01, HLA-A*11:01	En este estudio
MSYEDQDAL	4911-4919	I	HLA-B*35:01	En este estudio
NVIPTITQM	4926-4934	I	HLA-B*35:01	En este estudio

SAKNRARTV	4941-4949	I	HLA-B*08:01	En este estudio
MTNRQFHQK	4958-4966	I	HLA-A*11:01, HLA-A*03:01	En este estudio
KFYGGWHNMLK	4985-4995	I	HLA-A*11:01, HLA-A*23:01, HLA-A*24:02	En este estudio
NPHLMGWDYPK	5003-5011	I	HLA-A*03:01, HLA-A*11:01, HLA-B*35:01	En este estudio
MASLVLARK	5025-5033	I	HLA-A*11:01, HLA-A*03:01	En este estudio
CSLSHRFYRL	5038-5047	I	HLA-A*02:01, HLA-A*11:01, HLA-B*08:01	En este estudio
RLANCAQVL	5046-5054	I	HLA-A*02:01, HLA-B*35:01	En este estudio
SSGDATTAY	5073-5081	I	HLA-A*01:01, HLA-B*35:01	En este estudio
SVFNICQAV	5084-5092	I	HLA-A*02:01	En este estudio
LSDDAVVCF	5150-5158	I	HLA-A*01:01	En este estudio
SQGLVASIK	5164-5172	I	HLA-A*11:01	En este estudio
KQGDDYVYL	5213-5221	I	HLA-A*02:01	En este estudio
LMIERFVSLAIDAYPLTK	5246-5263	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-B*08:01, HLA-B*35:01, HLA-B*40:01	En este estudio
QEYADVFLHYL	5267-5277	I	HLA-A*01:01, HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01, HLA-B*40:01, HLA-B*44:02	En este estudio
HLYLQYIRK	5274-5282	I	HLA-A*03:01, HLA-A*11:01	En este estudio
WEPEFYEAMY	5308-5317	I	HLA-B*35:01, HLA-B*40:01	En este estudio
TPHTVLQAV	5318-5326	I	HLA-B*07:02	En este estudio
DVTDVTQLY	5380-5388	I	HLA-B*35:01	En este estudio
NGQVFGLYK	5410-5418	I	HLA-A*11:01	En este estudio
VLSRELHLSWEV	5481-5493	I	HLA-A*02:01 HLA-B*40:01, HLA-B*44:02, HLA-B*44:03	En este estudio
RPPLNRNYV	5497-5505	I	HLA-B*07:02	En este estudio
NYVFTGYRV	5503-5511	I	HLA-A*23:01, HLA-A*24:02	En este estudio
KVQIGEYTFEK	5516-5526	I	HLA-A*03:01, HLA-A*23:01	En este estudio
RITGLYPTL	5572-5580	I	HLA-A*02:01	En este estudio
AAVDALCEK	5636-5644	I	HLA-A*11:01	En este estudio
ALKYLPIDK	5645-5653	I	HLA-A*03:01, HLA-A*11:01	En este estudio
IPARARVEC	5658-5666	I	HLA-B*07:02	En este estudio
ALPETTADI	5686-5694	I	HLA-A*02:01	En este estudio
IVVFDEISM	5694-5702	I	HLA-B*35:01	En este estudio
VYIGDPAQL	5721-5729	I	HLA-A*23:01, HLA-A*24:02	En este estudio
LPAPRTLLT	5729-5737	I	HLA-B*07:02	En este estudio
CPAEIVDTV	5768-5776	I	HLA-B*35:01	En este estudio
GVVREFLTRNPA	5818-5829	I	HLA-A*11:01, HLA-B*40:01	En este estudio
LTRNPAWRKAVF	5824-5835	I	HLA-A*03:01, HLA-A*11:01, HLA-B*07:02, HLA-B*08:01, HLA-B*35:01	En este estudio
SEYDYVIFT	5863-5871	I	HLA-B*44:03	En este estudio
MSDRDLYDK	5900-5908	I	HLA-A*11:01	En este estudio
KLQFTSLEI	5908-5916	I	HLA-A*02:01	En este estudio
LQAENVTGL	5924-5932	I	HLA-A*02:01	En este estudio
GLCVDIPGIPK	5962-5972	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01	En este estudio
MTYRRLISM	5974-5982	I	HLA-B*07:02, HLA-B*08:01, HLA-B*35:01	En este estudio



YQVNGYPNM	5989-5997	I	HLA-A*02:01	En este estudio
EAVGTNLPL	6024-6032	I	HLA-B*35:01	En este estudio
VVRIKIVQM	6086-6094	I	HLA-B*08:01	En este estudio
NLSDRVVFVLW	6101-6111	I	HLA-A*02:01, HLA-B*44:02	En este estudio
TYACWHHSI	6148-6156	I	HLA-A*23:01, HLA-A*24:02	En este estudio
HSIGFDYVY	6154-6162	I	HLA-B*35:01	En este estudio
FMIDVQQWG	6165-6173	I	HLA-A*02:01	En este estudio
AVHECFVKR	6206-6214	I	HLA-A*11:01	En este estudio
VLHDIGNPK	6253-6261	I	HLA-A*03:01, HLA-A*11:01	En este estudio
QADVEWKFY	6268-6276	I	HLA-A*01:01	En este estudio
KFTDGVCLFW	6301-6309	I	HLA-A*01:01, HLA-A*23:01, HLA-A*24:02	En este estudio
KQLPFFYYS	6365-6373	I	HLA-A*02:01	En este estudio
QVVSIDIVY	6383-6391	I	HLA-A*02:01	En este estudio
HANEYRLYL	6412-6420	I	HLA-B*35:01	En este estudio
ISAGFSLWVYK	6427-6435	I	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-B*35:01	En este estudio
SLENVAFNVVNK	6453-6464	I	HLA-A*02:01, HLA-A*11:01, HLA-B*40:01	En este estudio
VSIINNTVYTK	6476-6486	I	HLA-A*03:01, HLA-A*11:01, HLA-B*35:01	En este estudio
KVDGVDVEL	6486-6494	I	HLA-A*02:01	En este estudio
TTLPVNVAFELWAKRNI	6499-6512	I	HLA-A*11:01, HLA-B*40:01	En este estudio
KPVPEVKIL	6516-6524	I	HLA-B*07:02	En este estudio
AANTVIWDY	6532-6540	I	HLA-B*35:01	En este estudio
CSMTDIACK	6554-6562	I	HLA-A*03:01, HLA-A*11:01	En este estudio
FRNARNGVL	6586-6594	I	HLA-B*08:01	En este estudio
VLITEGSVK	6593-6601	I	HLA-A*03:01, HLA-A*11:01	En este estudio
VTLIGEAVK	6617-6625	I	HLA-A*11:01	En este estudio
QSRNLQEFK	6648-6656	I	HLA-A*11:01	En este estudio
MEIDFLELAMDEFIERY	6661-6677	I	HLA-A*01:01, HLA-B*40:01, HLA-B*44:03	En este estudio
RYKLEGYAF	6676-6684	I	HLA-A*24:02, HLA-A*23:01	En este estudio
SQLGGLHLI	6695-6704	I	HLA-A*02:01, HLA-B*40:01	En este estudio
RFKESPFEL	6709-6717	I	HLA-A*23:01, HLA-A*24:02	En este estudio
FELEDFIPM	6715-6723	I	HLA-B*35:01, HLA-B*40:01	En este estudio
QAWQPGVAM	6801-6809	I	HLA-B*35:01	En este estudio
NLTLAVPY	6853-6861	I	HLA-B*35:01, HLA-A*11:01	En este estudio
PYNMRVIHF	6860-6868	I	HLA-A*24:02	En este estudio
CATVHTANK	6913-6921	I	HLA-A*11:01	En este estudio
DSKEGFFTY	6942-6950	I	HLA-B*35:01	En este estudio
YICGFIQKQ	6950-6958	I	HLA-A*11:01	En este estudio
ALGGSVAIK	6960-6968	I	HLA-A*03:01, HLA-A*11:01	En este estudio
KITEHSWNADLYK	6968-6980	I	HLA-A*02:01, HLA-A*03:01 HLA-A*11:01, HLA-B*40:01	En este estudio
KLMGHFAWWTAFV	6980-6992	I	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02	En este estudio
EAFILGCNYLG	7001-7011	I	HLA-A*02:01, HLA-B*35:01	En este estudio
REQIDGYVM	7014-7022	I	HLA-B*40:01	En este estudio
YVMHANYIFWRNTNPI	7020-7035	I	HLA-A*23:01, HLA-A*24:02, HLA-B*35:01	En este estudio

	PLKLRGTAV	7049-7057	I	HLA-B*08:01	En este estudio
ORF3a	LWLCWK	127-132	I	HLA-A*03:01	En este estudio
	YQIGGYTEK	184-192	I	HLA-A*11:01	En este estudio
	EEHVQIHTI	241-249	I	HLA-B*44:02, HLA-B*44:03	En este estudio
ORF6	LLIIMRTFK	15-23	I	HLA-A*03:01, HLA-A*11:01	En este estudio
	TFKVSINLDY	21-31	I	HLA-A*03:01, HLA-A*11:01, HLA-A*23:01	En este estudio
	SIWNLDYIINLI	25-36	I	HLA-A*02:01	En este estudio
S	MFVFLVLLPLVSSQCVNL T	1-19	II	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DPA1*02:02, HLA-DPB1*02:02	En este estudio
	TQSLIVNATNVVI	114-128	II	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01	En este estudio
	NVVIKVEFQFCN	125-137	II	HLA-DPA1*02:02, HLA-DPB1*02:02	En este estudio
	VADYSVLYNSASFST	362-376	II	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01	En este estudio
	DLCFTNVYADSFV	389-401	II	HLA-DRB1*07:01	En este estudio
	RQIAPGGTGKIA	408-419	II	HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	LQSYGFQPTNGVG	492-504	II	HLA-DRB1*04:01	En este estudio
	KKSTNLVKNKCV	528-539	II	HLA-DRB1*07:01	En este estudio
	TWRVYSTGSNVFQT	632-645	II	HLA-DRB1*04:01, HLA-DRB1*07:01	En este estudio
	GSNVFQTRAGCLIGAE	639-654	II	HLA-DRB1*07:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	FNFSQILPDPSPKPSK	800-814	II	HLA-DRB1*01:01, HLA-DRB1*04:01	En este estudio
	QKLIANQFNSAIGKIQDSL	920-938	II	HLA-DRB1*04:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	QLSSNFGAISSVLN	965-978	II	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	EVQIDRLITGRLQS	990-1003	II	HLA-DRB1*01:01	En este estudio
	IRASANLAATKMS	1018-1030	II	HLA-DRB1*04:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
N	PRITFGGSPDSTGSN	13-27	II	HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	EGSRGGSQASSR	174-185	II	HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	RQKRTATKAYNVT	259-271	II	HLA-DRB1*07:01	En este estudio
	WPQIAQFAPSASAFFGM	301-317	II	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
E	MYSFVSEETGLI	1-13	II	HLA-DRB1*04:01	En este estudio
	ETGLIVNSVLLFLAFVVF LL	8-28	II	HLA-DRB1*01:01, HLA-DRB1*07:01, HLA-DPA1*02:02,	En este estudio

	LLFLAFVVFLVTLAILTAL RLCAY	18-42	II	HLA-DPB1*02:02 HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPB1*02:01, HLA-DPB1*02:02	En este estudio
	IVNVSLVKPSFYVYS	46-60	II	HLA-DRB1*07:01, HLA-DPA1*02:02, HLA-DPB1*02:02	En este estudio
	PSFYVYSRVKLNLSRV PD	54-72	II	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01	En este estudio
ORF3a	YYQLYSTQLSTDT	211-223	II	HLA-DRB1*01:01, HLA-DRB1*04:01	En este estudio
	TIDGSSGVVNPVME	248-261	II	HLA-DPA1*02:02, HLA-DPB1*02:02, HLA-DQA1*05:01, HLA-DQB1*03:01, HLA-DQB1*03:01	En este estudio
	EPIYDEPTTTTSVPL	261-275	II	HLA-DRB1*04:01, HLA-DRB1*07:01	En este estudio
ORF6	DYIINLIKNLSKSLTEN	30-47	II	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01	En este estudio
ORF8	VAAFHQECSLQSCTQ	13-27	II	HLA-DRB1*04:01	En este estudio
ORF10	RNYIAQVDVVNF	24-35	II	HLA-DRB1*04:01	En este estudio
<b>Epítotos promiscuos</b>					
S	AYTNSFTRGVYY	27-38	I, II	HLA-A*11:01, HLA-B*35:01, HLA-DRB1*07:01	En este estudio
	RGVYYPDKVFRSSVLHS TQDLFLPFFSNVTFHAI HV	34-70	I, II	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01	En este estudio
	LPFNDGVYFASTEKSNI R	84-102	I, II	HLA-B*35:01, HLA-DRB1*07:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	NKSWMESEFRVYSSAN NCTFEYVSQPFLMDLE	149-180	I, II	HLA-A*01:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01	En este estudio
	KNLREFVFKNIDGYFKIY SKHTPI	187-210	I, II	HLA-A*11:01, HLA-A*23:01, HLA-DPA1*02:01, HLA-DPA1*02:02,	En este estudio

	<b>PIGINITRFQTLALHRSY LTPGDSSSGWTAGAAA YYVGYLQPRTEFLLYN</b>	230-280	I, II	HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01 HLA-A*01:01, HLA-A*02:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01, HLA-B*08:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	<b>KCTLKSFTVEKGIYQTSN FRVQPTEIVRFPNITNL CP</b>	300-337	I, II	HLA-A*03:01, HLA-A*11:01, HLA-B*35:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01	En este estudio
	<b>GEVFNATRFASVYAWN RK</b>	339-356	I, II	HLA-A*11:01, HLA-B*08:01, HLA-B*35:01, HLA-B*40:01, HLA-B*44:02, HLA-B*44:03, HLA-DRB1*07:01, HLA-DPA1*02:02, HLA-DPA1*02:01 HLA-DPB1*02:02, HLA-DPB1*02:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	<b>GNYNLYRLFRKSNLKP FERDISTEIYQAGSTPCN GVEGFNCY</b>	447-489	I, II	HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01, HLA-B*40:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	<b>QPYRVVLSFELLHAPA TVCG</b>	506-526	I, II	HLA-A*02:01, HLA-A*11:01, HLA-A*23:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01, HLA-DQA1*01:01, HLA-DQB1*05:01	En este estudio
	<b>CDIPIGAGICASYQTQTN SPRRARSVASQSIIAYTM S</b>	662-698	I, II	HLA-B*07:02, HLA-B*40:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DQA1*05:01,	En este estudio

				HLA-DQB1*03:01	
	<b>IPTNFTISVTTEILPV</b>	714-729	I, II	HLA-A*02:01, HLA-B*35:01, HLA-B*07:02, HLA-DRB1*07:01	En este estudio
	<b>AQKFNGLTVLPPLTDE MIAQYTSALLAGTITSGW TFGAGAALQIPFAM</b>	852-900	I, II	HLA-A*01:01, HLA-A*02:01, HLA-A*23:01, HLA-B*35:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	<b>GVVFLHVTYVPAQEKNF TTAPAICHD</b>	1059-1084	I, II	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DPA1*02:02, HLA-DPB1*02:02, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	<b>CDVVIGIVNNTVY</b>	1126-1138	I, II	HLA-B*35:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	<b>YEQYIKWPWYIWLGFIA GLIAIVMVTIM</b>	1206-1233	I, II	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*35:01, HLA-B*44:02, HLA-B*44:03, HLA-DRB1*01:01, HLA-DRB1*07:01, HLA-DQA1*05:01, HLA-DQB1*03:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPB1*02:01, HLA-DPB1*02:02	En este estudio
M	<b>GTITVEELKKLLEQWNL VIGFLFLTWCILLQFAYA NRNRFLYIIKLIFLWLLW PVTLACFV</b>	6-66	I, II	HLA-A*01:01, HLA-A*02:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-B*08:01, HLA-B*35:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DQA1*01:01, HLA-DQB1*05:01, HLA-DPA1*02:01, HLA-DPA1*02:02 HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01	En este estudio
	<b>AVILRGHLRIAGHHLGRC</b>	142-159	I, II	HLA-A*03:01, HLA-A*11:01, HLA-B*07:02, HLA-B*08:01, HLA-DRB1*01:01, HLA-DRB1*07:01	En este estudio
N	<b>SPDDQIGYRRATRRIR</b>	79-95	I, II	HLA-A*24:02, HLA-A*23:01, HLA-B*35:01, HLA-DRB1*07:01	En este estudio
	<b>LSPRWYFYLLGTGPEAG LPYGANKD</b>	104-128	I, II	HLA-A*01:01, HLA-A*11:01, HLA-B*07:02, HLA-B*08:01, HLA-DRB1*01:01, HLA-DQA1*05:01,	En este estudio

	<b>GTWLYTGAIKLDDK</b>	328-342	I, II	HLA-DQB1*03:01 HLA-A*23:01, HLA-A*24:02, HLA-DRB1*01:01, HLA-DRB1*07:01	En este estudio
	<b>QQTVTLLPAADLDDF</b>	389-403	I, II	HLA-B*35:01, HLA-DRB1*01:01	En este estudio
ORF3a	<b>MDLFMRIFTIGTVTLKQG</b>	1-18	I, II	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-B*08:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01	En este estudio
	<b>ASLPFGWLIVGVALLAVF QSASKIITLKK</b>	39-67	I, II	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*23:01, HLA-B*35:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DQB1*03:01, HLA-DQA1*05:01	En este estudio
	<b>LLLLFVTVYSHLLLVAAG LEAPFLYLYALVYFLQSI NFVRIIMRLWL</b>	83-129	I, II	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*07:02, HLA-B*08:01, HLA-B*35:01, HLA-B*44:02, HLA-B*44:03, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01, HLA-DQA1*05:01 HLA-DQB1*03:01	En este estudio
	<b>TGVEHVTFFIYNKIVDE</b>	223-239	I, II	HLA-A*03:01, HLA-A*11:01, HLA-B*44:02, HLA-B*44:03, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01	En este estudio
ORF7a	<b>MKIILFLALITLATCELYH Y</b>	1-20	I, II	HLA-A*01:01, HLA-A*02:01, HLA-A*11:01, HLA-B*35:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	<b>EVQELYSPIFLIVAAIVFIT LCFTLKRKT</b>	92-120	I, II	HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*23:01, HLA-A*24:02, HLA-B*07:02, HLA-B*35:01, HLA-B*40:01, HLA-B*44:03, HLA-DRB1*01:01, HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02,	En este estudio

				HLA-DPB1*23:01, HLA-DQA1*05:01, HLA-DQB1*03:01	
ORF7b	<b>MIELSLIDFYLCFLAFLLF LVLIMLIIFWF</b>	1-30	I, II	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*44:02, HLA-B*44:03, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01, HLA-DQA1*01:01, HLA-DQB1*05:01	En este estudio
ORF8	<b>MKFLVFLGIITTVAAFHQ</b>	1-18	I, II	HLA-A*02:01, HLA-A*23:01, HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	<b>PIHFYSKWYIRVGARKSA PLI</b>	38-58	I, II	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-DRB1*01:01, HLA-DRB1*04:01 HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01, HLA-DQA1*05:01, HLA-DQB1*03:01	En este estudio
	<b>IQYIDIGNYTVS</b>	71-82	I, II	HLA-A*02:01, HLA-DRB1*04:01	En este estudio
	<b>LVVRCSFYEDFLEYHDV RVV</b>	98-117	I, II	HLA-A*02:01, HLA-B*35:01, HLA-B*40:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01	En este estudio
ORF10	<b>MGYINVFAPFTIYSLLLC RMNSRNYI</b>	1-27	I, II	HLA-A*02:01, HLA-A*23:01, HLA-A*24:02, HLA-B*07:02, HLA-B*08:01, HLA-B*35:01, HLA-DRB1*01:01, HLA-DRB1*04:01 HLA-DRB1*07:01, HLA-DPA1*02:01, HLA-DPA1*02:02, HLA-DPA1*03:01, HLA-DPB1*02:01, HLA-DPB1*02:02, HLA-DPB1*23:01	En este estudio

**Tabla 9.** Epítomos reportados experimentalmente y capaces de ser reconocidos para montar una respuesta inmunitaria. En negrita se señalan los epítomos que se predijeron en este estudio. Anexo 1.

Proteína	Posición	Célula que reconoce	Secuencia	Fuente
<b>S</b>	562-580	Célula B lineal	<b>TESNKKFLPF</b> QQFGRDIA	(72)
<b>S</b>	819-835	Célula B lineal	PSKPSKRSFIEDLLFNKV	(72)
<b>IDa S</b>	21-45	Célula B lineal	TRTQLPPAYTNSFTRGVVYPDKVFRS	(73)
<b>IDb S</b>	221-245	Célula B lineal	SALEPLVDLPIGINITRFQTLALH	(73)
<b>IDc S</b>	261-285	Célula B lineal	GAAAYYVGYLQPRFTLLKYNENGTI	(73)
<b>IDd S</b>	330-349	Célula B lineal	PNITNLCPFGEVFNATRFAS	(73)
<b>IDe S</b>	375-394	Célula B lineal	STFKCYGVSPTKLNLCFTN	(73)
<b>IDf S</b>	450-469	Célula B lineal	NYLYRLFRKSNLKPFERDIS	(73)
<b>IDg S</b>	489-499	Célula B lineal	CNGVEGFNCYFPLQSYGFQP	(73)
<b>IDh S</b>	522-646	Célula B lineal	<b>ATVCGPKKSTNLVKNKCVNFNGLTGTGVLTE SNKKFLPF</b> QQFGRDIADTTDAVRDPQTLEILDITP CSFGGVS <b>VITPGTNTSNQVAVLYQDVNCTEVPV AIHADQLTPTWRVYSTGSNVFQTR</b>	(73)
<b>Idi S</b>	902-926	Célula B lineal	MAYRFNGIGVTQNVLYENQKLIANQ	(73)
<b>RBD S</b>	369-519	Célula B	YNSAFSTFKCYGVSPTKDLFDDFTNYLFAFNYQ YQTNGYFELH	(74)
<b>S</b>	16-52	Célula B lineal	VNLTRTQLPPAYTNSFTRGVVYPDKVFRSSVLH STQ	(75)
<b>RBD S</b>	439-478	Célula B lineal	NNLDSKVGGNLYLYRLFRKSNLKPFERDISTEY <b>QAGST</b>	(75)
<b>RBD S</b>	455-478	Célula B lineal	LFRKSNLKPFERDISTEY <b>QAGST</b>	(75)
<b>RBD S</b>	483-507	Célula B lineal	VEGFNCYFPLQSYGFQPT <b>NGVGYQP</b>	(75)
<b>S</b>	786-815	Célula B lineal	<b>KQIYKTPPIKDFGGFNFSQILPDP</b> SKPSKR	(75)
<b>S</b>	1238-1270	Célula B lineal	TSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKL	(75)
<b>M</b>	5-20	Célula B lineal	NGTITVEELKLLLEQW	(75)
<b>M</b>	198-208	Célula B lineal	RYRIGNYKLNLDHSSSSDNIA	(75)
<b>N</b>	153-170	Célula B lineal	NNAIVLQLPQGTTLPKG	(76)
<b>N</b>	121-138	Célula B lineal	LPYGANKDGIWVATEGA	(76)
<b>N</b>	129-146	Célula B lineal	GIIWVATEGALNTPKDHI	(76)
<b>N</b>	137-154	Célula B lineal	GALNTPKDHIGTRNPANN	(76)
<b>N</b>	145-162	Célula B lineal	HIGTRNPANNAIVLQLP	(76)
<b>S</b>	769-786	Célula B lineal	GI <b>AVEQDKNTQE</b> VFAQVK	(76)
<b>N</b>	101-120	Célula T CD4+	MKDLSRWFYFYLLGTGPEAG	(77)
<b>N</b>	81-95	Célula T CD4+	DDQIGYYRATRRIR	(77)
<b>N</b>	321-340	Célula T CD4+/CD8+	GMEVTPSGTWLTYTGAIKLD	(77)
<b>N</b>	104-121	Célula T CD4+/CD8+	LSPRWYFYLLGTGPEAGL	(78)
<b>S</b>	166-180	Célula T CD4+/CD8+	CTFEYVSQPFLMDLE	(78)
<b>S</b>	751-765	Célula T CD4+	NLLLQYGSFCTQLNR	(78)
<b>S</b>	801-815	Célula T CD4+	NFSQILPDPSKPSKR	(78)
<b>S</b>	866-880	Célula T CD4+	TDEMIAQYTSALLAG	(78)
<b>M</b>	141-158	Célula T CD4+	GAVILRGHLRIAGHHLGR	(78)
<b>M</b>	172-188	Célula T CD4+	TSRTLSYYKLGASQRVA	(78)
<b>M</b>	133-150	Célula T CD8+	LLESELVIGAVILRGHLR	(78)
<b>S</b>	1011-1028	Célula T CD4+	QLIRAAEIRASANLAATK	(78)
<b>S</b>	269-277	Célula T CD8+	YLQPRFTLL	(79), (80)
<b>S</b>	1020-1029	Célula T CD8+	ASANLAATK	(81), (82)
<b>S</b>	257-265	Célula T CD8+	WTAGAAAYY	(83), (84)
<b>S</b>	478-486	Célula T CD8+	TPCNGVEGF	(85), (86)
<b>N</b>	361-369	Célula T CD8+	KTFPPTEPK	(77), (87)



## Anexo 2. Identidad de epítomos de células B lineales y células T (%) compartida con otros coronavirus

En el anexo 2 se presentan las secuencias de los péptidos conservados con el porcentaje de identidad entre la secuencia del SARS-CoV-2 y los otros seis coronavirus de células B lineales (Tabla 10) y células T restringidos a MHC-I, MHC-II y promiscuos (Tabla 11). Las tablas están coloreadas según el porcentaje de identidad (más oscuro a mayor porcentaje de identidad).

**Tabla 10.** Epítomos conservados de células B con el porcentaje de identidad entre la secuencia de SARS-CoV-2 y seis coronavirus humanos. Anexo 2.

Proteína	SARS-CoV-2	SARS	MERS	229E	HKU1	NL63	OC43
S	DPFLGVYYHKNNKSWME	17.6	1.9	2.4	15.6	8.1	13.8
	MDLEGKQGNFKNL	53.8	0	13.3	0	13.3	0
	KHTPINLVRDLPGQFS	56.2	6.2	5.9	5.6	4.5	12.5
	GDEVQRQIAPGQTGKIADYNYKLPDD	92	20.7	2.4	20	5.9	23.1
	SNKKFLPF	62.5	0	50	7.7	25	7.7
	NCTEVPVAIHADQLTPT	64.7	3.6	6.2	23.5	3.1	29.4
	RVYSTGSNVFQ	81.8	0	9.1	9.1	9.1	18.2
	VNNSYECDIPI	81.8	21.4	6.7	12.5	6.7	15.4
	ASYQTQTNSPRRARSVASQ	31.6	40	6.5	24	15.8	15.8
	YTMSLGAENSVAYSNN	81.2	11.1	5	8	5	3.4
	KQIYKTPPIKDFGGF	80	40	20	3.2	13.3	15
	LADAGFIKQYGDCLG	86.7	35.3	20	46.7	18.8	46.7
	RNFYEQIITTD	83.3	41.7	41.7	25	16.7	33.3
	SGTNGTKRFDN	27.3	26.7	0	12.5	4.3	6.2
	NGTITD	100	33.3	9.1	50	9.1	33.3
	YQAGSTPCNGV	20	21.4	0	13.3	0	25
	YGFQPTNGVGYQ	66.7	5	8.3	10.5	6.7	5.3
	TVCGPKKSTN	80	20	8.3	17.6	20	23.1
	RDIADTTDAVRDPQ	64.3	4	4	4.2	0	8.3
	VEQDKNTQE	66.7	22.2	6.7	0	8.3	5.9
	ILPDPSPKSKRS	83.3	20	22.2	26.7	25	31.2
	DSLST	50	16.7	16.7	33.3	50	16.7
	PAQEKNTT	77.8	11.1	33.3	16.7	33.3	8.3
	KNHTSPDVLG	100	27.3	28.6	38.5	7.1	54.5
	VITPGTNTSN	80	9.1	10	20	9.1	20
	M	YRIGNYKLNTHSSSSDNIA	85	21.7	14.3	35	10
N	SKQLQQSMSSADS	100	0	5.6	4.2	5.6	5.9
	TGPEAGLPYGANK	92.3	69.2	23.1	46.2	30.8	46.2
	TDYKHW	100	33.3	11.1	16.7	33.3	14.3
	KLDDKDPNFKD	90	60	10	30	23.1	20
ORF1ab	DARTAPH	42.9	0	0	14.3	0	12.5
	QYGRSGE	85.7	8.3	0	11.1	0	7.7
	GNKGAGGHSYGA	91.7	33.3	0	0	0	41.7
	DELGTDPYEDFQENWNTHSSG	77.3	10	5.4	7.1	0	7.1
	GGAYTR	83.3	50	42.9	0	14.3	14.3
	FCGPDGYPL	100	55.6	55.6	8.3	55.6	8.3
	ERSEKSYELQ	80	40	30	27.3	50	8.3
	PRVEKK	100	33.3	22.2	33.3	14.3	0
	PVASEPNEC	87.5	6.7	12.5	10	11.1	12.5
	CGETSWQTG	66.7	66.7	55.6	33.3	44.4	33.3
	NLTKEGATTCGYL	76.9	4	4.5	14.3	4.5	23.5
	HNSEVGPEHSLAEYHNES	61.1	21.7	0	4.8	0	4
	GVVGESEGL	40	14.3	5.6	25	6.7	18.8
	VTKGKAKKGAWNIG	85.7	42.9	28.6	4.3	21.4	0
	LQPLEQPTSEAVEAPLV	29.4	26.3	9.5	17.4	4.5	29.4
	TLKGGAPTQVTFG	78.6	71.4	35.7	30.8	35.7	9.1
	YPPDEDEEEGDCEEEEFEPST	67.5	19.3	17.3	6.4	23.1	28.9
	QYEGTDDYQGKPLEFG						

TSAALQPEEEQEEDWLDDDSQQTVGQQDG SEDNQTTT	29.7	17.9	0	5.8	0	25.6
IVVEAKVKV	44.4	22.2	8.3	44.4	0	44.4
GGVAGALNKATNNAMQVESDDYIATNGPLK V	83.9	58.1	38.7	29	54.8	25.8
PNVKNKGE	71.4	14.3	25	11.1	28.6	22.2
EKQVEQ	28.6	0	28.6	0	0	0
AEIPKEEVKPFITESKPSVEQRKQDDKK	12.8	7.1	12.5	0	18.8	0
NLHPDSAT	37.5	10	0	12.5	0	11.1
TKKAGGTTE	88.9	0	0	0	0	0
KVPTDNYITTPGQGLNGYTVEE	78.3	8.7	0	0	0	0
TVSVSSPDAVTA	83.3	7.1	11.1	0	16.7	0
LTSSSKTPEEH	90.9	12.5	0	0	0	0
YKDWSYSYGQSTQ	75	23.1	0	25	0	16.7
KSVYYTSPNPTTFH	50	42.9	5.9	14.3	10	3.8
GQQFGPTYLDGADVTKIKPHNSHEG	92	28	20	36	28.1	36
YHTTDPFS	75	50	25	7.1	0	7.1
KWKYPQVN	75	37.5	25	44.4	25	44.4
NPPALQD	71.4	57.1	0	28.6	8.3	0
YRARAGE	85.7	14.3	14.3	57.1	14.3	28.6
GELGDV	100	33.3	50	33.3	50	22.2
GQQQTTL	71.4	28.6	0	28.6	0	22.2
APPAQYEL	75	12.5	0	25	0	25
EYTGNYQ	100	8.3	0	0	0	14.3
SSEYKGPI	75	25	0	18.2	12.5	12.5
NSYTTTIK	87.5	12.5	0	9.1	0	20
DPKLDNYYKKDNSYFTE	76.5	52.9	0	35.3	0	35.3
IDLVPNQYPNASF	85.7	13.6	5	8.3	13.6	8.3
LTGYKKPASR	70	33.3	0	0	0	0
HYTPSFK	71.4	42.9	8.3	0	8.3	14.3
ATNKATYKP	66.7	18.2	0	9.1	0	20
KPVETSN	85.7	42.9	0	14.3	0	12.5
SEDAQGM	71.4	11.1	0	18.2	0	10
LKPVSEEVVENPTIQ	80	15	6.7	11.8	6.7	4.5
NSRIKAS	85.7	28.6	28.6	14.3	14.3	14.3
PTTIKKN	100	10	10	14.3	7.1	14.3
SLDTYPSL	75	50	25	50	18.2	37.5
PINPTDQSSY	100	70	20	70	27.3	50
AGQKTYE	100	57.1	9.1	42.9	9.1	42.9
ANNTKGS	100	18.2	7.1	11.1	7.1	11.1
KSKCEESSAKSA	75	41.7	16.7	11.8	25	11.8
DVGDSAE	85.7	71.4	9.1	0	9.1	22.2
GFVSDSVETK	70	60	23.1	60	23.1	60
DIEVTGDSC	88.9	44.8	7.1	22.2	7.1	44.4
VENMTPRD	100	12.5	22.2	25	14.3	25
IDGGVTRDIAS	72.7	36.4	18.2	63.6	18.2	45.5
SQRGGSYTNDK	90.9	33.3	16.7	23.1	25	30.8
DASGKVPYCY	90	60	40	50	30	60
TCERSE	100	33.3	33.3	50	33.3	33.3
AMDTTSYR	87.5	62.5	25	62.5	25	62.5
DFSNSGSDVLYQPPQTSI	94.4	61.1	50	57.9	38.9	57.9
FPSGKV	100	66.7	66.7	33.3	66.7	50
DTANPKTPKY	90	70	30	60	40	50
NGSPSGV	100	57.1	42.9	57.1	57.1	57.1
GVHAGTDLEGN	90.9	36.4	36.4	63.6	45.5	54.5
VDRQTAQAAGTDTT	100	35.7	14.3	28.6	14.3	28.6
VYDDGAR	85.7	14.3	9.1	0	0	0
GLLPPKNSI	88.9	44.4	33.3	66.7	33.3	77.8
FATAQEAYEQAVANGDS	94.1	41.2	58.8	20.8	52.9	20.8
SEFDRD	100	16.7	66.7	50	66.7	50
KQARSEDKRAKV	100	75	41.7	58.3	50	58.3
YNTYKNTCDGTTT	84.6	15.8	21.4	9.5	31.2	9.5
SMDNSPNL	87.5	37.5	25	25	25	12.5
AAGTTQACTDDN	100	38.5	11.8	11.8	5	30.8
FPKSDGTGTIYTELEP	100	29.4	31.2	37.5	31.2	37.5
RFVTDTPKGP	100	45.5	72.7	36.4	54.5	36.4

	GNATEVPAN	100	44.4	44.4	66.7	44.4	55.6
	LASGGQPIT	100	44.4	22.2	44.4	22.2	44.4
	TGTGQAITVTPEANMDQESFG	100	61.9	47.6	57.1	52.4	57.1
	DHPNPKGF	100	37.5	50	37.5	37.5	37.5
	TTCANDPVG	100	55.6	33.3	22.2	33.3	33.3
	QEKDEDDNL	77.8	22.2	22.2	44.4	22.2	33.3
	IQTTPGSGVPVV	58.3	58.3	33.3	66.7	33.3	33.3
	SHVDTDLTK	66.7	44.4	10	11.1	20	11.1
	DNTSRYWEP	100	44.4	22.2	20	11.1	6.2
	KSHKPPI	100	28.6	42.9	42.9	42.9	42.9
	CVGSDNVT	100	37.5	25	37.5	37.5	37.5
	TFEKGDYGDA	100	60	54.5	20	45.5	20
	KAHKDKS	100	42.9	28.6	57.1	9.1	42.9
	PGIPKD	100	16.7	12.5	9.1	12.5	28.6
	CHATREAVG	88.9	55.6	44.4	66.7	22.2	44.4
	TGYVDTPNNTDFSRVSAKPPPGDQF	84	44	34.4	36	44	40
	AQPCSDKAY	100	44.4	11.1	7.1	22.2	33.3
	THSDKFT	85.7	14.3	28.6	42.9	28.6	42.9
	DSPCESHGKQ	100	50	36.4	36.4	30	30
	HFDGQQGEV	66.7	55.6	44.4	55.6	44.4	55.6
	KRDAPAH	85.7	0	14.3	0	14.3	8.3
	IAKKPTET	87.5	8.3	12.5	25	12.5	18.2
	VKGLQPSVGPQASL	73.3	33.3	16	33.3	22.2	46.7
	LQEFKPRSQ	77.8	33.3	66.7	55.6	55.6	22.2
	AQTGSS	100	12.5	33.3	50	14.3	33.3
	KLQSSQAWQPGVA	92.3	53.8	38.5	38.5	46.2	38.5
	QNYGDSATL	55.6	33.3	44.4	44.4	44.4	55.6
	YDPKTKNVTKENDSKE	87.5	68.8	56.2	56.2	43.8	56.2
	GKPREQI	85.7	42.9	0	42.9	8.3	42.9
ORF3a	QGEIKDATPSDFVR	18.8	-	-	-	-	-
	KIITLKKRWQL	81.8	-	-	-	-	-
	SKNPLL	100	-	-	-	-	-
	PYNSVT	100	-	-	-	-	-
	TSGDGTTSPISEHDYQIGGYTEKWESGV	57.1	-	-	-	-	-
	STQLSTDTGV	70	-	-	-	-	-
	VDEPEEHVQIHTIDGSSGVVNPVME PIYDEPTTTTS	68	-	-	-	-	-
	EPIYDEPTTTTSVP	92.9	-	-	-	-	-
ORF6	LTENKYSQLDEEQP	57.1	-	-	-	-	-
ORF7a	LYHYQECVR	90	-	-	-	-	-
	EPCSSGTYEGNSPFHPLAD	94.7	-	-	-	-	-
	VKHVYQLRARSVSPKLFIRQEEVQEL	85.2	-	-	-	-	-
ORF7b	ELQDHNE	57.1	-	-	-	-	-
ORF8	QSCTQHQPYYVDDPCPIHFYSKW	2.9	13	-	-	-	-
	RVGARKSAP	11.1	11.1	-	-	-	-
	DEAGSKSPIQYIDIGN	0	12.5	-	-	-	-
	EDFLEY	0	0	-	-	-	-

**Tabla 11. Epítomos conservados de células T con restricción MHC-I, MHC-II y promiscuos con el porcentaje de identidad entre la secuencia de SARS-CoV-2 y seis coronavirus humanos. Anexo 2.**

	Proteína	SARS-CoV-2	SARS	MERS	229E	HKU1	NL63	OC43
MHC-I	S	YQDVNCTEV	88.9	22.2	7.1	33.3	0	22.2
		RFDNPVLPF	77.8	15.4	0	16.7	25	25
		FPREGVFVS	88.9	33.3	6.2	33.3	6.2	44.4
		LEPLVDLPI	44.4	22.2	18.2	5.9	18.2	5.9
		FQFCNDPFL	44.4	15.4	5.9	22.2	6.2	22.2
		NVYADSFVI	88.9	7.1	6.2	33.3	7.7	33.3
		TLADAGFIK	88.9	44.4	11.1	44.4	11.1	44.4
		QMAYRFNGI	100	44.4	22.2	44.4	22.2	44.4
		NQFNSAIGK	66.7	55.6	33.3	33.3	44.4	44.4
		AQALNTLVK	100	55.6	33.3	66.7	44.4	55.6
		RAAEIRASANL	100	36.4	45.5	18.2	36.4	11.1
		FVSNQTHWFVTQRNF	73.3	30	17.6	26.7	27.8	26.7
		NASVVNIQK	88.9	11.1	8.7	22.2	8.7	22.2
	M	SSDNIALLV	88.9	33.3	30	44.4	30	44.4
		RYRIGNYKL	100	55.6	22.2	44.4	22.2	44.4
	N	KHWPQIAQF	100	55.6	44.4	22.2	44.4	22.2
		LPNNTASWF	100	66.7	8.3	27.3	6.7	33.3
		RIRGGDGKM	88.9	33.3	33.3	20	30	20
		ASAFFGMSR	100	77.8	20	30	30	66.7
		GMSRIGMEV	100	33.3	0	7.7	15.4	7.7
		KTFPPTPEK	100	66.7	8.3	7.1	16.7	0
	ORF1ab	SLVPGFNEK	77.8	8.3	0	22.2	0	9.1
		VLQVRDVLV	100	9.1	0	5.9	0	5.9
		VLSEARQHLK	80	7.1	0	5.6	0	18.2
		GLVEVEKGV	88.9	11.5	0	10.5	0	13
		RTAPHGHVM	33.3	22.2	11.1	8.3	11.1	6.7
		VMVELVAEL	66.7	13.3	7.1	18.2	7.1	13.3
		AELEGIQYG	77.8	18.2	7.7	0	7.7	5.9
		TLGVLVPHV	100	33.3	7.1	44.4	33.3	22.2
		VLVPHVGEIPVAY	84.6	15.4	7.7	25	17.4	12.5
		YPLECIKDL	77.8	22.2	6.7	5.9	6.7	5.9
		TLSEQLDFI	88.9	0	8.3	11.1	22.2	11.1
		REHEHEIAW	88.9	11.1	33.3	22.2	22.2	22.2
		YELQTPFEI	88.9	44.4	33.3	44.4	44.4	33.3
		KTIQPRVEK	88.9	33.3	13.3	6.2	18.2	0
		KLDGFMGRIRSVY	84.6	13.6	13.3	33.3	13.3	4.2
		CEFCGTENL	88.9	30	7.7	30	33.3	33.3
		KEGATTCGY	77.8	6.7	7.7	6.7	7.7	6.7
		SEVGPEHSL	66.7	18.2	0	10	0	22.2
		TIAFGGCVF	76.8	44.4	6.2	33.3	22.2	55.6
		FSYVGCHNK	77.8	33.3	6.2	11.1	7.7	10
		SEGLNDNLLEI	63.6	6.7	0	4.3	7.7	4.3
		NINIVGDFK	88.9	44.4	11.1	8.3	0	8.3
		KLNEEIAII	77.8	22.2	10	13.3	4.8	0
		STSAFVETVK	70	7.1	30	10	20	0
		GAWNIGEYK	77.8	33.3	6.2	0	0	0
		ILSPYAFA	44.4	12.5	11.8	15.4	10.5	16.7
AARVVRSIFSR		72.7	4	5.6	14.3	10.5	6.7	
YSLRLIDAM		77.8	33.3	5.9	6.7	44.4	33.3	
LATNNLVVM		44.4	10	15.4	18.2	15.4	0	
VMAYITGGVVQL		66.7	18.8	11.1	13.6	7.4	7.7	
WLTNIFGTVEYKL		61.5	5.9	38.5	23.1	13.6	5.9	
KLKPVLDWL		44.4	22.2	11.1	22.2	0	11.1	
FLRDGWEIV		66.7	33.3	20	0	5	9.1	
CEIVGGQIV	55.6	6.2	6.7	6.2	0	13.3		
ALCADSIII	33.3	0	11.8	0	5.9	6.2		
ALNLGETFV	66.7	5.9	9.5	20	10.5	9.1		
FLEGETLPTEVL	41.7	7.1	4.8	35.7	22.2	16.7		

TEVLTEEVVLK	63.6	16.7	45.5	23.1	18.2	11.1
SEAVEAPLV	7.7	18.2	33.3	6.7	6.7	33.3
TPVCINGLM	88.9	22.2	6.7	33.3	6.7	22.2
KYCALAPNM	55.6	7.1	9.1	6.7	22.2	7.1
NMMVTNNTFTLK	50	33.3	21.4	8.3	6.7	0
TLKGGAPTK	77.8	77.8	15.4	0	30	9.1
FGDDTVIEVQGYK	84.6	46.2	31.2	21.4	28.6	3.8
YTVELGTEV	88.9	11.1	0	44.4	16.7	15.4
CVVADAVIK	77.8	14.3	11.1	44.4	10	44.4
SELLTPLGIDL	72.7	28.6	0	4.5	0	9.5
YELQTPFEI	88.9	44.4	33.3	44.4	44.4	33.3
KTIQPRVEK	88.9	33.3	13.3	6.2	18.2	0
KLDGFMGRIRSVY	84.6	13.6	13.3	33.3	13.3	4.2
CEFCGTENL	88.9	30	7.7	30	33.3	33.3
KEGATTCGY	77.8	6.7	7.7	6.7	7.7	6.7
SEVGPEHSL	66.7	18.2	0	10	0	22.2
TIAFGGCVF	76.8	44.4	6.2	33.3	22.2	55.6
FSYVGCHNK	77.8	33.3	6.2	11.1	7.7	10
SEGLNDNLLEI	63.6	6.7	0	4.3	7.7	4.3
NINIVGDFK	88.9	44.4	11.1	8.3	0	8.3
KLNEEIAII	77.8	22.2	10	13.3	4.8	0
STSAFVETVK	70	7.1	30	10	20	0
GAWNIGEQQK	77.8	33.3	6.7	0	6.7	0
ILSPYAFA	44.4	12.5	11.8	15.4	10.5	16.7
AARVVRSIFSR	72.7	4	5.6	14.3	10.5	6.7
YSLRLIDAM	77.8	33.3	5.9	6.7	44.4	33.3
LATNNLVVM	44.4	10	15.4	18.2	15.4	0
VMAYITGGVVQL	66.7	18.8	11.1	13.6	7.4	7.7
WLTNIFGTVEK	61.5	5.9	38.5	23.1	13.6	5.9
KLKPVLDWL	44.4	22.2	11.1	22.2	0	11.1
FLRDGWEIV	66.7	33.3	20	0	5	9.1
CEIVGGQIV	55.6	6.2	6.7	6.2	0	13.3
ALCADSIII	33.3	0	11.8	0	5.9	6.2
ALNLGETFV	66.7	5.9	9.5	20	10.5	9.1
FLEGETLPTEVL	41.7	7.1	4.8	35.7	22.2	16.7
TEVLTEEVVLK	63.6	16.7	45.5	23.1	18.2	11.1
SEAVEAPLV	7.7	18.2	33.3	6.7	6.7	33.3
TPVCINGLM	88.9	22.2	6.7	33.3	6.7	22.2
KYCALAPNM	55.6	7.1	9.1	6.7	22.2	7.1
NMMVTNNTFTLK	50	33.3	21.4	8.3	6.7	0
TLKGGAPTK	77.8	77.8	15.4	0	30	9.1
FGDDTVIEVQGYK	84.6	46.2	31.2	21.4	28.6	3.8
YTVELGTEV	88.9	11.1	0	44.4	16.7	15.4
CVVADAVIK	77.8	14.3	11.1	44.4	10	44.4
SELLTPLGIDL	72.7	28.6	0	4.5	0	9.5
YLFDESGEF	66.7	33.3	44.4	66.7	44.4	66.7
TTIQTIVEV	0	33.3	0	10	0	11.1
LEMELTPVV	22.2	11.8	5.7	4.9	12.1	4.9
KLTDNVYIK	88.9	33.3	36.4	33.3	23.1	33.3
KPTVVVNAANVYLK	64.3	57.1	42.9	42.9	50	50
YIATNGPLK	66.7	55.6	33.3	11.1	77.8	11.1
VLSGHNLAK	88.9	66.7	0	44.4	0	44.4
CLHVVGPNVVK	81.8	54.5	36.4	36.4	36.4	45.5
YENFNQHEV	55.6	22.2	20	22.2	14.3	33.3
HEVLLAPLL	66.7	22.2	33.3	22.2	44.4	11.1
SLRVCVDTV	77.8	7.7	44.4	33.3	16.7	33.3
NVYLAVFDKNL	63.6	18.2	5.3	27.3	16.7	7.7
KIKACVEEV	77.8	0	0	0	0	0
VTTTLEETK	100	0	0	22.2	0	22.2
KFLTENLLLY	70	10	9.1	21.4	10	30
LYIDINGNL	66.7	22.2	11.1	6.7	0	6.2
LVSDIDITFLK	27.3	0	6.2	18.2	18.2	7.7
APYIVGDVV	77.8	11.1	11.1	8.3	7.7	7.1
QEGVLTAVV	44.4	0	0	0	0	0
VLTAVVIPTKK	63.6	0	0	0	0	0

PTDNYITTY	77.8	0	0	0	0	0
SAFYILPSI	77.8	0	0	0	0	0
QEILGTVSW	88.9	0	0	0	0	0
KLMPVCVETKA	54.5	27.3	0	9.1	0	9.1
KAIVSTIQR	66.7	7.7	11.1	27.3	11.1	9.1
VVDYGARFY	66.7	40	7.1	21.4	14.3	33.3
SLINTLNDLNETL	61.5	13.3	17.6	13.3	4.8	13.3
MPLGYVTHG	85.7	85.7	0	0	9.1	0
DAVTAYNGY	88.9	9.1	6.7	0	8.3	0
TQLGIEFLK	77.8	13.3	0	14.3	0	21.4
GEVITFDNLK	60	10	16.7	7.7	7.7	20
SLREVRTIKVFTTV	90.9	23.1	36.4	18.2	5.3	5.6
YGQQFGPTY	100	33.3	33.3	22.2	33.3	22.2
LPNDDTLRV	77.8	25	8.3	8.3	22.2	16.7
YYHTTDPSTF	77.8	44.4	22.2	13.3	6.7	13.3
YMSALNHTK	100	18.2	10	0	5.9	20
YPQVNGLTS	77.8	8.3	33.3	30	7.7	36.4
YLATALLTL	55.6	44.4	11.1	44.4	20	22.2
ELKFNPAL	77.8	55.6	7.7	44.4	7.1	6.2
ALQDAYYRA	88.9	22.2	22.2	33.3	8.3	33.3
ALILAYCNK	88.9	66.7	10	41.7	0	41.7
SYLFQHANL	66.7	14.3	30.8	22.2	15.4	22.2
CASEYTGNY	88.9	6.7	33.3	22.2	33.3	22.2
TLYCIDGALLTK	81.8	18.2	0	16.7	0	28.6
VVCTEIDPK	66.7	44.4	0	22.2	0	33.3
YPNASFDNFKFV	75	11.8	4.5	41.7	11.8	25
FPDLNGDVV	100	44.4	6.7	55.6	44.4	55.6
TYKPNTWCI	77.8	0	0	0	0	0
CIRCLWSTK	88.9	11.1	0	10	0	7.7
VVENPTIQK	100	33.3	11.1	16.7	11.1	9.1
TEVVGDIILK	53.3	4	13.3	18.8	4.5	11.8
TEEVGHTDLM	70	30	0	10	0	10
VPWDTIANY	55.6	0	0	22.2	0	22.2
STTTNIVTR	44.4	33.3	0	11.1	0	11.1
NYMPYFFTLLL	81.8	18.2	35.7	9.1	28.6	36.4
ASMPPTIAK	88.9	13.3	10.5	11.1	10.5	11.1
FCLEASFNYL	50	20	10	30	7.7	5.6
LINIIWFLLLSV	53.8	23.1	5.3	7.7	23.1	6.2
STNVTIATY	44.4	11.1	22.2	10	7.1	10
YPSLETIQI	77.8	44.4	11.1	22.2	8.3	22.2
SFKWDLTAF	55.6	7.7	5.9	9.1	5.9	12.5
TAFGLVAEWF	60	5.9	5.6	20	5.6	20
GLVAEWFLAYILF	76.9	33.3	5	30.8	11.1	30.8
AYILFTRFFYVL	75	50	5.6	50	14.3	50
FYVLGLAAIM	80	40	27.3	21.4	14.3	45.5
IMQLFFSYFAVHFI	78.6	14.3	5.9	28.6	21.4	9.1
WLMWLIINLVQM	75	9.5	5.6	26.7	5.6	26.7
AMVRMYIFFASFYYVWK	94.1	35.3	0	29.4	30	35.3
YKRNRATRV	88.9	88.9	8.3	66.7	44.4	77.8
TTIVNGVRR	77.8	44.4	77.8	66.7	44.4	55.6
SYIVDSVTV	88.9	77.8	30	22.2	33.3	16.7
GSIHLYFDK	77.8	9.1	33.3	22.2	33.3	33.3
SLPINVIVF	100	30	33.3	18.2	33.3	30
KMFDAYVNTF	90	70	20	70	30	70
FSSTFNVPM	77.8	44.4	7.7	44.4	7.1	44.4
ATAEAELAK	77.8	22.2	22.2	33.3	22.2	33.3
VLSTFISAAR	90	80	0	50	0	70
QVVNVVTTK	88.9	11.1	22.2	33.3	22.2	44.4
KVTLVFLFV	33.3	30	22.2	15.4	30	20
TLVFLVAAIFYL	30.8	30.8	46.2	17.6	35.7	0
FVAAIFYLITPVHVM	33.3	33.3	8	11.8	15.8	11.8
FSSEIIGYKAI	72.7	18.2	38.5	20	30.8	20
STDTCFANK	88.9	66.7	22.2	66.7	22.2	66.7
REVGFFVPLPGTIL	80	26.7	20	20	33.3	26.7
KLIEYTDFA	88.9	33.3	11.1	44.4	11.1	33.3

LAAECTIFK	100	55.6	22.2	55.6	22.2	44.4
NVLEGSVAY	66.7	55.6	44.4	22.2	44.4	11.1
SLRPDTRYVLM	90.9	45.5	18.2	63.6	27.3	54.5
SIIQFPNTY	100	30	20	30	20	30
TYLEGSVRVV	100	20	5.3	50	0	60
VVTTFDSEY	88.9	22.2	33.3	33.3	22.2	44.4
SLPGVFCGVDAV	75	50	25	58.3	25	66.7
MFTPLIQPI	66.7	44.4	16.7	55.6	6.2	44.4
IVAGGIVAI	77.8	40	5.9	44.4	7.1	44.4
IVVTCLAYYFMRFR	78.6	28.6	20	21.4	21.4	21.4
YFMRFRRAFG	80	40	50	50	40	50
VVAFNTLLFL	80	12.5	20	30	23.1	30
LMSFTVLCL	88.9	11.1	16.7	23.1	16.7	15.4
LPGVYSVIY	88.9	27.3	22.2	33.3	22.2	44.4
VYSVIYLYLTFYLT	85.7	50	21.4	42.9	21.4	42.9
VPFWITIAIICI	69.2	38.5	23.1	30.8	15.4	30.8
ISTKHFWFFSNYL	71.4	29.4	5	33.3	4.3	33.3
FEEAALCTFLLNK	100	46.2	53.8	53.8	53.8	61.5
TQYNRYLAL	100	44.4	22.2	66.7	22.2	55.6
MLNPNYEDL	100	44.4	10	44.4	0	55.6
QLRVIGHSM	100	66.7	44.4	33.3	44.4	33.3
GTDLEGNFY	88.9	33.3	33.3	77.8	44.4	66.7
FLNRFTTTLNDFNLVA	100	31.2	4.3	31.2	25	31.2
LLLTLTSL	80	20	0	36.4	8.3	14.3
STQWSLFFFL	90	30	13.3	20	13.3	20
FFFLYENAFLPF	91.7	21.4	35.7	23.1	14.3	25
FLPFAMGIIAMSA	61.5	9.5	30.8	4.5	9.1	13.3
FLCLFLLPSLATVAYFNMV	100	52.6	47.4	31.6	47.4	31.6
VYMPASWVM	100	23.1	7.7	6.2	6.7	6.2
WVMRIMTWLDMV	75	12.5	5	5	4.5	21.4
YASAVLLI	88.9	33.3	15.4	6.2	15.4	18.2
ILMTARTVYDDGARRVWTL	94.7	9.5	5.6	16.7	6.5	12.5
SMWALIISV	88.9	16.7	11.1	20	22.2	7.1
YSGVVTTVMFLARGIVF	88.2	23.5	9.1	27.8	13.6	23.5
FMCVEYCIPIFITGNTL	76.5	13.6	23.8	11.5	21.1	8
MLVYCFLGY	100	44.4	44.4	55.6	44.4	55.6
YFCTCYFGL	77.8	66.7	50	44.4	33.3	55.6
CYFGLFCLLNRYFRLTLGV	100	52.6	47.4	57.9	42.1	63.2
QLHNDILLA	100	66.7	44.4	55.6	55.6	55.6
LAKDTTEAF	100	55.6	33.3	33.3	33.3	33.3
MQGAVDINK	88.9	33.3	5.9	20	6.7	25
LPSYAAFAT	88.9	25	44.4	11.1	44.4	11.1
SEFDRDAAM	100	22.2	44.4	44.4	44.4	44.4
ALNNIINNA	100	66.7	33.3	66.7	22.2	66.7
IPLTTAAKL	100	50	60	40	60	40
KLMVVIPDYNTYK	84.6	53.8	38.5	30.8	46.2	30.8
TCDGTTFTY	88.9	14.3	27.3	33.3	30	33.3
TYASALWEI	100	66.7	33.3	55.6	33.3	55.6
SPNLAWPLI	100	55.6	55.6	33.3	55.6	33.3
CTDDNALAYNNTTK	85.7	28.6	35.7	20	50	28.6
KSDGTGTIY	100	22.2	22.2	33.3	6.7	33.3
CAFAVDAAK	88.9	44.4	66.7	66.7	66.7	66.7
GTSTDVVYR	100	77.8	44.4	66.7	44.4	66.7
VAGFAKFLK	100	55.6	50	33.3	44.4	55.6
NLIDSYFVVK	90	70	30	60	50	50
HEETIYNLL	88.9	44.4	66.7	44.4	66.7	22.2
RVYANLGERV	100	70	70	30	70	40
SGVPVDSY	77.8	77.8	44.4	66.7	55.6	55.6
MPILTLTRA	100	44.4	33.3	44.4	55.6	55.6
AESHVDTDL	77.8	55.6	22.2	22.2	33.3	22.2
KPYIKWDLKY	90.9	54.5	45.5	36.4	45.5	36.4
VLVSTVFP	100	44.4	44.4	55.6	44.4	55.6
LLVYAADPAM	100	80	40	90	40	80
SVAALTNNVAF	100	63.6	45.5	54.5	46.7	63.6
QLLFVVEVDKY	100	66.7	25	83.3	25	83.3

		MSYEDQDAL	100	66.7	66.7	33.3	77.8	44.4
		MASLVLARK	100	77.8	22.2	77.8	33.3	77.8
		SQGLVASIK	88.9	33.3	33.3	44.4	33.3	44.4
		KQGDDYVYL	100	55.6	18.2	66.7	27.3	60
		QEYADVFHLYL	100	54.5	45.5	54.5	45.5	54.5
		HLYLQYIRK	100	66.7	11.1	55.6	11.1	55.6
		WEPEFYEAMY	100	50	50	50	50	40
		TPHTVLQAV	100	55.6	55.6	33.3	44.4	33.3
		VLSDRELHLSWEV	100	38.5	38.5	53.8	53.8	61.5
		NYVFTGYRV	100	77.8	44.4	77.8	44.4	77.8
		RITGLYPTL	100	77.8	33.3	18.2	33.3	18.2
		ALKYLPIDK	100	66.7	33.3	44.4	44.4	44.4
		GVVREFLTRNPA	100	50	33.3	33.3	41.7	33.3
		MSDRDLYDK	100	33.3	66.7	33.3	44.4	33.3
		KLQFTSLEI	100	44.4	22.2	33.3	36.4	55.6
		LQAENVTGL	100	66.7	44.4	5.6	33.3	20
		GLCVDIPGIPK	100	33.3	25	5.6	25	14.3
		MTYRRLISM	100	44.4	44.4	66.7	44.4	66.7
		YQVNGYPNM	100	44.4	6.7	22.2	11.1	22.2
		TYACWHSI	77.8	66.7	33.3	55.6	33.3	55.6
		AVHECFVKR	100	55.6	66.7	44.4	77.8	44.4
		QADVEWKFY	88.9	22.2	33.3	55.6	33.3	44.4
		QVVSIDIDYV	100	50	22.2	38.5	22.2	33.3
		ISAGFSLWVYK	90.9	54.5	45.5	63.6	45.5	63.6
		VSIINNTVYTK	90.9	36.4	27.3	36.4	27.3	54.5
		KVDGVDVEL	77.8	30	30	33.3	16.7	44.4
		KPVPEVKIL	88.9	33.3	33.3	55.6	33.3	33.3
		AANTVIWDY	100	44.4	44.4	44.4	44.4	55.6
		CSMTDIAKK	88.9	44.4	44.4	36.4	33.3	36.4
		FRNARNGVL	100	22.2	44.4	77.8	44.4	44.4
		VLITEGSVK	100	22.2	44.4	22.2	33.3	44.4
		VTLIGEAVK	88.9	6.7	5.9	27.3	7.1	18.2
		QSRNLQEFK	66.7	44.4	66.7	55.6	55.6	44.4
		MEIDFLELAMDEFIERY	88.2	52.9	47.1	35.3	52.9	52.9
		RYKLEGYAF	100	66.7	44.4	66.7	44.4	66.7
		RFKESPFEL	44.4	9.1	5.9	33.3	5.9	9.1
		FELEDIFPM	77.8	13.3	22.2	11.1	22.2	11.1
		QAWQPGVAM	100	55.6	33.3	44.4	44.4	55.6
		NLTLAVPY	100	77.8	55.6	77.8	44.4	77.8
		CATVHTANK	100	33.3	55.6	23.1	55.6	6.2
		YICGFIQQK	77.8	33.3	66.7	55.6	44.4	55.6
		KLMGHFAWWTAFV	92.3	69.2	38.5	61.5	46.2	69.2
		REQIDGYVM	77.8	55.6	27.3	55.6	14.3	55.6
		PKLRGTAV	100	66.7	33.3	77.8	33.3	88.9
	ORF3a	IMRLWLCWK	88.9	-	-	-	-	-
		YQIGGYTEK	77.8	-	-	-	-	-
		EEHVQIHTI	66.7	-	-	-	-	-
	ORF6	LLIIMRTFK	77.8	-	-	-	-	-
		TFKVSINWLDY	63.6	-	-	-	-	-
		SIWNLDYIINLI	66.7	-	-	-	-	-
MHC-II	S	MFVFLVLLPLVSSQCVNLT	34.8	20.6	0	19.2	21.1	13.9
		RQIAPGQTGKIA	91.7	16.7	5	16.7	5.3	13.3
		LQSYGFQPTNGVG	53.8	23.1	5	15	5.3	10
		KKSTNLVKNKCV	66.7	28.6	33.3	25	41.7	33.3
		TWRVYSTGSNVFQT	78.6	9.5	0	7.1	5	14.3
		TQSLLVNATNVVI	66.7	4.2	4.7	29.4	2.4	11.4
		NVVIKVEFQFCN	53.8	0	13.3	20	0	23.1
		VADYSVLVNSASFST	93.3	17.6	0	8.3	0	9.5
		DLCFTNVYADSFV	92.3	23.1	14.3	30.8	0	30.8
		GSNVFQTRAGCLIGAE	87.5	33.3	4.3	25	8.7	31.2
		FNFSQILPDPSKPSK	86.7	20	25	20	25	35
		QKLIANQFNSAIGKIQDSL	73.7	57.9	42.1	52.6	47.4	57.9
		IRASANLAATKMS	100	38.5	46.2	30.8	38.5	30.8
		PRITFGGSDSTGSN	80	26.7	4	6.7	4.5	13.3
		EGSRGGSQASSR	100	66.7	15.4	58.3	10	58.3



		RQKRTATKAYNVT	92.3	53.8	40	61.5	46.2	46.2
		WPQIAQFAPSASAFFGM	100	76.5	38.9	44.4	38.9	50
	E	MYSFVSEETGLI	100	46.2	5	15.4	15.4	4.3
		ETGTLIVNSVLLFLAFVVFLL	100	28.6	33.3	28.6	28.6	3.4
		LLFLAFVVFLLVTLAILTALRLCAY	100	40	32	36	20	33.3
		IVNVSLVKPSFYVYS	86.8	40	26.7	40	13.3	33.3
		PSFYVYSRVKLNLSRVPD	80	21.7	4.5	39.1	2.9	30.4
	ORF3a	YYQLYSTQLSTDT	76.9	-	-	-	-	-
		TIDGSSGVVNPVME	78.6	-	-	-	-	-
		EPIYDEPTTTTSVPL	93.3	-	-	-	-	-
	ORF6	DYIINLIKNSKSLTEN	44.4	-	-	-	-	-
	ORF8	VAAFHQECSLQSQCTQ	6.7	0	-	-	-	-
	ORF10	RNYIAQVDVNVF	2.2	-	-	-	-	-
MHC-I & MHC-II	S	AYTNSFTRGVYY	46.7	23.1	0	28.6	17.6	31.2
		RGVYYDPKVFRRSSVLHSTQDLFLP	70.3	25	0	18.8	7.3	10.5
		FFSNVTWFHAIHV						
		LPFNDGVYFASTEKSNIIR	68.4	13.9	0	17.2	11.1	17.9
		NKSWMESEFRVYSSANNCTFEYV	31.2	14.3	16.7	19.5	8	11.1
		KNLREFVFKNIDGYFKIYSKHTPI	58.3	7.1	16.7	9.7	16.7	25
		PIGINITRFQTLAL	46.7	3.7	9.4	13.6	18.8	3.7
		NITRFQTLALHRSYLTPGDSSSG	45.8	24.6	5.5	24	8.8	27.7
		WTAGAAAYYGYLQPRTFLLKYN						
		KCTLKSFTVEKGIYQTSNFRVQPT	36.8	17.1	15.9	16.9	14.6	16.9
		GEVFNATRFASVYAWNPK	83.3	11.5	15	27.8	15.8	31.6
		GNYNLYRLFRKSNLKPFERDISTE	47.6	16.4	15.9	5.2	14.8	12.1
		IYQAGSTPCNGVEGFNCY						
		QPYRVVLSFELLHAPATVCG	95.2	2.3	2.4	2.9	7.5	3.1
		CDIPIGAGICASYQTQNSPRRARS	62.2	34.1	17.1	27.9	14.6	22.2
		IPTNFTISVTTEILPV	75	50	50	50	50	37.5
		AQKFNGLTVLPPLTDEMIAQYTS	85.7	46.9	38.8	38.8	38.8	36.7
	LLAGTITSGWTFGAGAALQIPFAM							
	GVVFLHVTYVPAQEKNFTTAPAICH	88.5	30.8	38.5	30.8	38.5	30.8	
	CDVVIGIVNNTVY	92.3	4.8	15.4	35.3	15.4	17.6	
	YEQYIKWPWYIWLGFIAGLIAIMV	92.9	60.7	42.9	40.6	36.4	44.8	
	M	GTITVEELKLLLEQWNLVIGFLFT	90.2	29.5	30.8	34.3	29.5	39.3
		WICLLQFAYANRNRFLYIIKLIFLWL						
		AVILRGHLRIAGHHLGRC	83.3	55.6	33.3	50	27.8	50
	LPKEITVATSRTLSYYKLGASQRVA	91.4	40	36.1	34.3	38.9	42.9	
	N	SPDDQIGYYRRATRRIR	88.2	41.2	29.4	27.8	41.2	27.8
		LSPRWYFYLLGTGPEAGLPYGAN	92	76	39.4	64	48	68
		GTWLTYTGAIKLDDK	93.3	60	20	37.5	20	40
		QQTVTLLPAADLDDF	86.7	20	4.2	16.7	26.7	15.8
	ORF3a	MDLFMRIFTIGTVTLKQG	50	-	-	-	-	-
		ASLPFGWLIVGVALLAVFQSASKIIT	79.3	-	-	-	-	-
		LALSKGVHFVCNLLLLFVTYVSHLL	79.7	-	-	-	-	-
		LVAAGLEAPFLYLYALVYFLQSINF						
	TGVEHVTFIYNKIVDE	58.8	-	-	-	-	-	
	ORF7a	MKILFLALITLATCELYHY	75	-	-	-	-	-
		EVQELYSPIFLIVAAIVFITLCFTLKR	80	-	-	-	-	-
	ORF7b	MIELSLIDFYLCFLAFLFLVLIIMLIIF	93.3	-	-	-	-	-
	ORF8	MKFLVFLGIITVAAFHQ	0	0	-	-	-	-
		PIHFYSKWYIRVGARKSAPLI	18.2	23.8	-	-	-	-
		IQYIDIGNYTVS	4.8	16.7	-	-	-	-
		LVVRCSFYEDFLEYHDRVV	21.1	0	-	-	-	-
	ORF10	MGYINVFAPFTIYSLLLCRMNSRN	7.5	-	-	-	-	-

### Anexo 3. Portada del *pre-print* del manuscrito elaborado y depositado en el servidor *bioRxiv*

En el anexo 3 se presenta la primera hoja del manuscrito realizado en conjunto con la elaboración de esta tesis (Figura 12), en el cual se integran las variantes de emergencia, incluyendo la variante Omicron.

bioRxiv preprint doi: <https://doi.org/10.1101/2022.02.16.480759>; this version posted February 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

#### **Bioinformatic analysis of B and T cell epitopes from SARS-CoV-2 Spike, Membrane and Nucleocapsid proteins as a strategy to assess possible cross-reactivity between emerging variants, including Omicron, and other human coronaviruses**

1 **Diana Laura Pacheco-Olvera<sup>1§</sup>, Stephanie Saint Remy-Hernández<sup>1§</sup>, María**  
2 **Guadalupe García-Valeriano<sup>1</sup>, Tania Rivera-Hernández<sup>1,2\*</sup>, Constantino López-**  
3 **Macías<sup>1\*</sup>**

4 <sup>1</sup>Unidad de Investigación Médica en Inmunoquímica, UMAE Hospital de Especialidades del Centro  
5 Médico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, México

6 <sup>2</sup> CONACYT, México

7

8 <sup>§</sup> **These authors contributed equally**

9 **\* Correspondence:**

10 Tania Rivera-Hernández or Constantino López-Macías

11 [trivera@conacyt.mx](mailto:trivera@conacyt.mx) or [constantino@sminmunologia.mx](mailto:constantino@sminmunologia.mx), [constantino.lopez@imss.gob.mx](mailto:constantino.lopez@imss.gob.mx)

12

13

#### 14 **Abstract (342)**

15 The COVID-19 pandemic caused by SARS-CoV-2 produced a global health emergency  
16 since December 2019, that up to the end of January 2022 had caused the death of more than  
17 5.6 million people worldwide. Despite emergence of new variants of concern, vaccination  
18 remains one of the most important tools to control the pandemic. All approved vaccines and  
19 most of the vaccine candidates use the spike protein of the virus as a target antigen to induce  
20 protective immune responses. Several variants of the virus present key mutations in this  
21 protein which render the virus, at different rates, to evade the neutralizing antibody response.  
22 Although experimental evidence suggests that cross-reactive responses between  
23 coronaviruses are present in the population, it is unknown which potential antigens shared  
24 between different coronaviruses could be responsible for these responses. This study  
25 provides predictions of new potential B and T cell epitopes within SARS-CoV-2 Spike (S),  
26 Membrane (M) and Nucleocapsid (N) proteins together with a review of the reported B  
27 epitopes of these proteins. We also analyse amino acid changes present in the epitopes of  
28 variants of concern (VOC) and variants being monitored (VBM), and how these might affect  
29 the immune response, as these changes may alter the peptides' immunogenicity index and

**Figura 12. *Pre-print* del manuscrito**