

VACUNAS DE ARN CONTRA EL CORONAVIRUS

SOLICITUD RELACIONADA

Esta solicitud reivindica el beneficio según el Título 35 del Código de los Estados Unidos, sección 119, literal (e) de la solicitud provisional estadounidense número 62/967,006, presentada el 28 de enero de 2020, la solicitud provisional estadounidense número 62/971,825, presentada el 7 de febrero de 2020, la solicitud provisional estadounidense número 63/002,094, presentada el 30 de marzo de 2020, la solicitud provisional estadounidense número 63/009,005, presentada el 13 de abril de 2020 y la solicitud provisional estadounidense número 63/016,175, presentada el 27 de abril de 2020, cada una de las cuales se incorpora en su totalidad a la presente mediante referencia.

ANTECEDENTES

Los coronavirus humanos son virus de ARN monocatenario, positivos, con envoltura y altamente contagiosos de la familia *Coronaviridae*. Se conoce que dos subfamilias de *Coronaviridae* causan enfermedades humanas. Las más importantes son los β -coronavirus (beta-coronavirus). Los β -coronavirus son agentes etiológicos comunes de infecciones leves a moderadas del tracto respiratorio superior. Sin embargo, los brotes de nuevas infecciones por coronavirus, como las infecciones causadas por un coronavirus de Wuhan, se han asociado con una alta tasa de mortalidad. Un coronavirus 2 de síndrome respiratorio agudo grave [SRAS] (SARSCoV-2, según su sigla en inglés) (anteriormente denominado "coronavirus de Wuhan", "nuevo coronavirus de 2019" o "2019-nCoV") se identificó inicialmente en la ciudad china de Wuhan en diciembre de 2019 y ha infectado rápidamente a cientos de miles de personas. La enfermedad pandémica que causa el virus SARSCoV-2 ha sido nombrado por la Organización Mundial de la Salud (OMS) como COVID-19 (Enfermedad por coronavirus 2019). La primera secuencia del genoma de un aislado de SARS-CoV-2 (también conocido como 2019 nCoV o Wuhan-Hu-1) fue depositada en GenBank el 12 de enero de 2020 por investigadores del CDC (Centro Chino para el Control y la Prevención de Enfermedades) en Pekín.

Actualmente, no existe un tratamiento específico para la COVID-19 o una vacuna para la infección por SARS-CoV-2. Los continuos problemas de salud y la mortalidad asociados con las infecciones por coronavirus, en particular la pandemia del SARS-CoV-2, son una gran

preocupación a nivel internacional. La crisis de salud pública causada por el SARS-CoV-2 refuerza la importancia de desarrollar rápidamente opciones de vacunas eficaces y seguras contra estos virus.

5 **COMPENDIO**

En el presente documento, en algunas modalidades, se proporcionan composiciones inmunizantes (p. ej., vacunas de ARN) que comprenden un ARN que codifica antígenos altamente inmunogénicos capaces de provocar potentes respuestas de anticuerpos neutralizantes contra antígenos de coronavirus, tales como antígenos del SARS-CoV-2.

10 Sorprendentemente, las secuencias de antígenos proteicos de este nuevo coronavirus comparten menos del 80 % de identidad con las secuencias de antígenos proteicos del coronavirus del síndrome respiratorio agudo grave (SARS), y menos del 35 % de identidad con las secuencias de antígenos proteicos del coronavirus del síndrome respiratorio de Oriente Medio (MERS, por sus siglas en inglés).

15 Las construcciones proporcionadas en el presente documento, en algunas modalidades, incluyen: una reversión del sitio de escisión polibásico en la proteína espiga (S) del SARS-CoV-2 nativo a un único sitio de escisión básico (p. ej., Figura 1, Variante 7, SEQ ID NO: 23); una eliminación de la secuencia señal polibásica ER/Golgi (KXHXX-COOH) en la cola carboxi (p. ej., Figura 1, Variante 8, SEQ ID NO: 26); una mutación estabilizadora de prolina

20 doble (p. ej., Figura 1, Variantes 1-6 y 9, SEQ ID NO: 5, 8, 11, 14, 17, 20 y 29); un sitio de escisión de proteasa modificado para estabilizar la proteína (p. ej., Figura 1, Variantes 3 y 5, SEQ ID NO: 11 y 17); una eliminación de la cola citoplásmica (p. ej., Figura 1, variantes 3, 4 y 6, SEQ ID NO: 11, 14 y 20); y/o un armazón de pliegue (p. ej., Figura 1, Variantes 3 y 4, SEQ ID NO: 11 y 14). Las características estructurales descritas en la presente incluyen,

25 por ejemplo, la supresión del sitio de escisión de furina al sustituirlo opcionalmente por una región transmembrana, un pliegue injertado en la porción C-terminal del ectodominio del pico, la cola intracelular C-terminal eliminada (cola carboxi). Por lo tanto, el ARNm proporcionado en la presente, en algunas modalidades, comprende un marco de lectura abierto que codifica una variante de proteína de pico trimérica que comprende cualquiera o

30 más de un sitio de escisión de furina eliminado, secuencia adicional de pliegue como extremo C, cola o secuencia de carboxi eliminada en este y/o mutación de 2 prolina.

Algunos aspectos de la presente descripción proporcionan un ácido ribonucleico (ARN) que comprende un marco de lectura abierto (ORF) que codifica un antígeno de coronavirus (p. ej., una proteína S, una proteína de membrana (M), una proteína de envoltura (E), una proteína de nucleocápside (NC), o una proteína de la Tabla 1) capaz de inducir una
 5 respuesta inmune (p. ej., una respuesta de anticuerpo neutralizante) al SARS-CoV-2, opcionalmente en donde el ARN se formula en una nanopartícula lipídica.

Otros aspectos de la presente descripción proporcionan un ARN de codón optimizado que comprende un ORF que comprende una secuencia que tiene al menos un 80 % de identidad con un ARN de tipo salvaje que codifica un antígeno del SARS-CoV-2, opcionalmente en el
 10 que el ARN se formula en una nanopartícula lipídica.

Incluso otros aspectos de la presente descripción proporcionan un ARN modificado químicamente que comprende un ORF que comprende una secuencia que tiene al menos un 80 % de identidad con un ARN de tipo salvaje que codifica un antígeno del SARS-CoV-2, opcionalmente en donde el ARN se formula en un lípido nanopartícula.

15 Incluso otros aspectos de la presente descripción proporcionan un ARN que comprende un ORF que comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de cualquiera de las secuencias de la Tabla 1, p. ej., SEQ ID NO: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 u 84. Incluso En algunas modalidades, el ARN comprende un ORF que comprende una secuencia
 20 que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 28. En algunas modalidades, el ARN comprende un ORF que comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 16. En algunas modalidades, el ARN comprende un ORF que comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 19. En algunas modalidades, el ARN comprende un ORF
 25 que comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 22. En algunas modalidades, el ARN comprende un ORF que comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 25.

En algunas modalidades, el ORF comprende una secuencia que tiene al menos 85 %, al
 30 menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de cualquiera de las secuencias de la Tabla 1, p. ej., SEQ ID NO: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 u 84.. En algunas modalidades,

el ARN comprende un ORF que comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 28. En algunas modalidades, el ARN comprende un ORF que comprende la secuencia de SEQ ID NO: 28. En algunas modalidades, el ARN comprende un ORF que comprende una

5 secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 16. En algunas modalidades, el ARN comprende un ORF que comprende la secuencia de SEQ ID NO: 16. En algunas modalidades, el ARN comprende un ORF que comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la

10 secuencia de SEQ ID NO: 19. En algunas modalidades, el ARN comprende un ORF que comprende la secuencia de SEQ ID NO: 19. En algunas modalidades, el ARN comprende un ORF que comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 22. En algunas modalidades, el ARN comprende un ORF que comprende la secuencia de SEQ ID

15 NO: 22. En algunas modalidades, el ARN comprende un ORF que comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 25. En algunas modalidades, el ARN comprende un ORF que comprende la secuencia de SEQ ID NO: 25. En algunas modalidades, el ARN comprende un ORF que comprende una secuencia que tiene al menos

20 el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 106. En algunas modalidades, el ARN comprende un ORF que comprende la secuencia de SEQ ID NO: 106. En algunas modalidades, los ARNm que comprenden los ORF se modifican de manera uniforme (*p. ej.*, se modifican completamente, se modifican en toda la secuencia) para una modificación particular. Por ejemplo, un ARN

25 puede modificarse de manera uniforme con 1-metil-pseudouridina, de manera que cada U en la secuencia sea una 1-metil-pseudouridina.

En algunas modalidades, el ARN comprende además una UTR 5' [región no traducida], en donde opcionalmente la UTR 5' comprende la secuencia de SEQ ID NO: 2 o SEQ ID NO:

30 36.

En algunas modalidades, el ARN comprende además una UTR 3', en donde opcionalmente la UTR 3' comprende la secuencia de SEQ ID NO: 4 o SEQ ID NO: 37.

En algunas modalidades, el ARN comprende además un análogo de cubierta de 5', opcionalmente una cubierta de 7mG (5') ppp (5') NImpNp. Pueden utilizarse otros análogos de cubierta.

En algunas modalidades, el ARN comprende además una cola poli (A), que opcionalmente tiene una longitud de 50 a 150 nucleótidos.

En algunas modalidades, el ORF codifica un antígeno de coronavirus. En algunas modalidades, el antígeno de coronavirus es una proteína estructural. En algunas modalidades, la proteína estructural es una proteína de pico (S). En algunas modalidades, la proteína S es una forma de prefusión estabilizada de una proteína S. En algunas

modalidades, el antígeno del coronavirus comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de cualquiera de las secuencias de la Tabla 1, p. ej., SEQ ID NO: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83 u 85. En algunas modalidades, el antígeno de coronavirus comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO:

29. En algunas modalidades, el antígeno de coronavirus comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 17. En algunas modalidades, el antígeno de coronavirus comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 20. En algunas modalidades, el antígeno de coronavirus comprende una secuencia que tiene al menos un 80 % de identidad con la

secuencia de SEQ ID NO: 23. En algunas modalidades, el antígeno de coronavirus comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 26. En algunas modalidades, el antígeno de coronavirus comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de cualquiera de las secuencias de la Tabla 1, p. ej.,

SEQ ID NO : 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83 u 85. En algunas modalidades, el antígeno del coronavirus comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 29. En algunas modalidades, el

antígeno de coronavirus comprende la secuencia de SEQ ID NO: 29. En algunas modalidades, el antígeno de coronavirus comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 17. En algunas modalidades, el antígeno de coronavirus comprende la

secuencia de SEQ ID NO: 17. En algunas modalidades, el antígeno del coronavirus comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 20. En algunas modalidades, el antígeno de coronavirus comprende la secuencia de SEQ ID NO: 20. En algunas modalidades, el antígeno de coronavirus comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 23. En algunas modalidades, el antígeno de coronavirus comprende la secuencia de SEQ ID NO: 23. En algunas modalidades, el antígeno del coronavirus comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 26. En algunas modalidades, el antígeno de coronavirus comprende la secuencia de SEQ ID NO: 26.

En algunas modalidades, la proteína estructural es una proteína M. En algunas modalidades, la proteína M comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 81. En algunas modalidades, la proteína M comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 81. En algunas modalidades, el ORF comprende la secuencia de SEQ ID NO: 80. En algunas modalidades, el ARN comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 95. En algunas modalidades, el ARN comprende la secuencia de SEQ ID NO: 95.

En algunas modalidades, la proteína estructural es una proteína E. En algunas modalidades, la proteína E comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 83. En algunas modalidades, la proteína E comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 83. En algunas modalidades, el ORF comprende la secuencia de SEQ ID NO: 82. En algunas modalidades, el ARN comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 96. En algunas modalidades, el ARN comprende la secuencia de SEQ ID NO: 96.

En algunas modalidades, la proteína estructural es una proteína NC. En algunas modalidades, la proteína NC comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 85. En algunas modalidades, la proteína NC

comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 85. En algunas modalidades, el ORF comprende la secuencia de SEQ ID NO: 84. En algunas modalidades, el ARN comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 97. En algunas modalidades, el ARN comprende la secuencia de SEQ ID NO: 97.

En algunas modalidades, el ORF comprende la secuencia de cualquiera de las secuencias de la Tabla 1, p. ej., cualquiera de las SEQ ID NO: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84 o 106. En algunas modalidades, los ARNm que comprenden los ORF se modifican de manera uniforme (p. ej., se modifican completamente, se modifican en toda la secuencia) para una modificación particular. Por ejemplo, un ARN puede modificarse de manera uniforme con 1-metil-pseudouridina, de manera que cada U en la secuencia sea una 1-metil-pseudouridina.

En algunas modalidades, el ARN comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 %, o al menos 98 % de identidad con la secuencia de cualquiera de las secuencias de la Tabla 1, p. ej., cualquiera de SEQ ID NO: 1, 6, 9, 12, 15, 18, 21, 24, 27, 30, 51, 53, 55, 57, 58, 60, 86-97 o 105. En algunas modalidades, el ARN comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 27. En algunas modalidades, el ARN comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 105. En algunas modalidades, el ARN comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 15. En algunas modalidades, el ARN comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 18. En algunas modalidades, el ARN comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 21. En algunas modalidades, el ARN comprende una secuencia que tiene al menos 85 %, al menos 90 %, al menos 95 % o al menos 98 % de identidad con la secuencia de SEQ ID NO: 24.

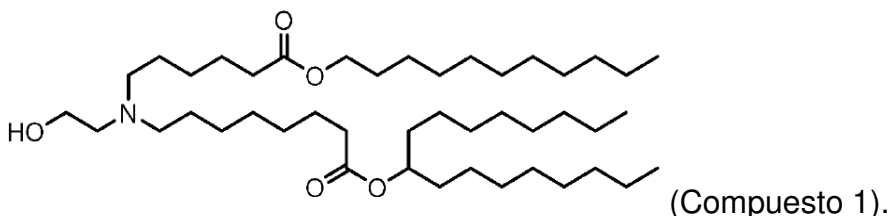
En algunas modalidades, el ARN comprende la secuencia de cualquiera de las secuencias

En algunas modalidades, el ARN comprende una modificación química. En algunas modalidades, la modificación química es 1-metilpseudouridina (p. ej., completamente modificada, modificada a lo largo de toda la secuencia).

Algunos aspectos de la presente descripción proporcionan un método que comprende un codón que optimiza el ARN de cualquiera de las modalidades anteriores.

En algunas modalidades, el ARN se formula en una nanopartícula lipídica.

En algunas modalidades, la nanopartícula lipídica comprende un lípido modificado con PEG, un lípido no catiónico, un esteroil, un lípido catiónico ionizable o cualquier combinación de estos. En algunas modalidades, la nanopartícula lipídica comprende 0.5-15 % en moles (p. ej., 0.5-10 % en moles, 0.5-5 % en moles o 1-2 % en moles) de lípidos modificados con PEG; 5-25 % en moles (p. ej., 5-20 % en moles, o 5-15 % en moles) *de lípidos no catiónicos* (p. ej., neutros); 25-55 % en moles (p. ej., 30-45 % en moles o 35-40 % en moles) de esteroil; y 20-60 % en moles (p. ej., 40-60 % en moles, 40-50 % en moles, 45-55 % en moles o 45-50 % en moles) de lípido catiónico ionizable. En algunas modalidades, el lípido modificado con PEG es 1,2 dimiristoil-sn-glicerol, metoxipolietilenglicol (PEG2000 DMG), el lípido no catiónico es 1,2 diestearoil-sn-glicerol-3-fosfolina (DSPC), el esteroil es colesterol; y el lípido catiónico ionizable tiene la estructura del Compuesto 1:



Otros aspectos de la presente descripción proporcionan una composición que comprende el ARN de cualquiera de las modalidades anteriores y una mezcla de lípidos. En algunas modalidades, la mezcla de lípidos comprende un lípido modificado con PEG, un lípido no catiónico, un esteroil, un lípido catiónico ionizable o cualquier combinación de los mismos.

- 5 En algunas modalidades, la mezcla de lípidos comprende 0.5-15 % en moles (p. ej., 0.5-10 % en moles, 0.5-5 % en moles o 1-2 % en moles) de lípidos modificados con PEG; 5-25 % en moles (p. ej., 5-20 % en moles, o 5-15 % en moles) *de lípidos no catiónicos* (p. ej., neutros); 25-55 % en moles (p. ej., 30-45 % en moles o 35-40 % en moles) de esteroil; y 20-60 % en moles (p. ej., 40-60 % en moles, 40-50 % en moles, 45-55 % en moles o 45-50 %
10 en moles) de lípido catiónico ionizable. En algunas modalidades, el lípido modificado con PEG es 1,2 dimiristoil-sn-glicerol, metoxipolietilenglicol (PEG2000 DMG), el lípido no catiónico es 1,2 diestearoil-sn-glicerol-3-fosfocolina (DSPC), el esteroil es colesterol; y el lípido catiónico ionizable tiene la estructura del Compuesto 1.

- En algunas modalidades, la mezcla de lípidos forma nanopartículas lipídicas. En algunas
15 modalidades, el ARN se formula en las nanopartículas lipídicas. En algunas modalidades, las nanopartículas lipídicas se forman primero como nanopartículas lipídicas vacías y se combinan con el ARNm de la vacuna inmediatamente antes (p. ej., dentro de un par de minutos hasta una hora de) la administración.

- Otros aspectos adicionales de la presente descripción proporcionan un método que
20 comprende administrar a un sujeto el ARN de cualquiera de las modalidades anteriores en una cantidad eficaz para inducir una respuesta de anticuerpos neutralizantes contra un coronavirus en el sujeto.

- Incluso otros aspectos adicionales de la presente descripción proporcionan un método que
25 comprende administrar a un sujeto la composición de cualquiera de las modalidades anteriores en una cantidad eficaz para inducir una respuesta de anticuerpos neutralizantes y/o una respuesta inmune de células T, opcionalmente un CD4⁺ y/o una respuesta inmune de células T CD8⁺ contra un coronavirus en el sujeto.

En algunas modalidades, el coronavirus es un SARS-CoV-2.

- En algunas modalidades, el sujeto está inmunodeprimido. En algunas modalidades, el sujeto
30 tiene una enfermedad pulmonar. En algunas modalidades, el sujeto tiene 5 años de edad o menos, o 65 años de edad o más.

En algunas modalidades, el método comprende administrar al sujeto al menos dos dosis de

la composición.

En algunas modalidades, se producen niveles detectables del antígeno de coronavirus en el suero del sujeto entre 1 y 72 horas luego de la administración del ARN o composición que comprende el ARN.

- 5 En algunas modalidades, se produce un título de anticuerpo neutralizante de al menos 100 NU/ml, al menos 500 NU/ml o al menos 1000 NU/ml en el suero del sujeto entre 1 y 72 horas luego de la administración del ARN o composición que comprende el ARN.

Debe entenderse que los términos “SARS-CoV-2”, “coronavirus de Wuhan”, “nuevo coronavirus de 2019” y “2019-nCoV” se refieren al mismo betacoronavirus de reciente
10 aparición ahora conocido como SARS-CoV-2 y se utilizan indistintamente en este documento.

Todo el contenido de la Solicitud Internacional núm. PCT/US2016/058327 (Publicación núm. WO2017/07062) y la Solicitud Internacional núm. PCT / US2018/022777 (Publicación núm. WO2018/170347) se incorporan aquí como referencia.

15

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La **Figura 1** muestra esquemas de varios antígenos ilustrativos de la proteína S codificados por el ARNm de SARS-CoV-2 de la presente descripción. El esquema superior representa una proteína SARS-CoV-2 de tipo salvaje; los esquemas siguientes representan variantes
20 de la proteína SARS-CoV-2, en relación con el tipo salvaje.

La **Figura 2** muestra un gráfico de datos de expresión *in vitro* de 24 horas para diversas variantes de la proteína SARS-CoV-2 codificadas por el ARNm de SARS-CoV-2 de la presente descripción.

La **Figura 3** muestra un gráfico de datos de expresión *in vitro* de 24 horas para diversas
25 variantes de la proteína SARS-CoV-2 codificadas por el ARNm de SARS-CoV-2 de la presente descripción. Se probaron dos cantidades diferentes de ARNm.

Las **Figuras 4A-4B** muestran gráficos de las mediciones del título de anticuerpos en suero luego de la inmunización con diferentes dosis de la vacuna de ARNm de SARS-CoV-2 de la variante 9 en diferentes cepas de ratones (Figura 4A) y en dosis más altas (Figura 4B).

30 Las **Figuras 5A-5C** muestran gráficos de las mediciones del título de anticuerpos séricos luego de la inmunización con diferentes dosis de la vacuna de ARNm de la variante 5 del SARS-CoV-2 (Figura 5A), en comparación con la vacuna de ARNm de la variante 9 del

SARS-CoV-2 y un ARNm que codifica una proteína S de SARS-CoV-2 de tipo salvaje (Figura 5B). La Figura 5C es un gráfico que compara el título de anticuerpos en suero para siete vacunas de ARNm de SARS-CoV-2 diferentes y un ARNm que codifica la secuencia de proteína S de SARS-CoV-2 de tipo salvaje.

- 5 La **Figura 6** muestra un gráfico de una respuesta temporal de anticuerpos en ratones luego de la inmunización con el ARNm de la variante 9 del SARS-CoV-2 a diferentes dosis.

La **Figura 7** muestra un esquema que muestra los programas de dosificación.

- Las **Figuras 8A-8C** muestran gráficos de los títulos de anticuerpos séricos en ratones dos semanas luego de una dosis inicial de la vacuna de ARNm de la variante 9 del SARS-CoV-2 y dos semanas luego de una dosis de refuerzo de la Vacuna de ARNm de la variante 9 de Wuahn-Hu-1 en ratones BALB/c (Figura 8A), ratones C57BL / 6 (Figura 8B) y ratones C3B6 (Figura 8C). Se probaron varias dosis de vacuna.

- Las **Figuras 9A-9E** muestran gráficos de los títulos de anticuerpos séricos de ratones dos semanas luego de una dosis inicial de la vacuna de ARNm de la variante 5 del SARS-CoV-2 y dos semanas luego de una dosis de refuerzo de la variante 5 del SARS-CoV-2 Vacuna de ARNm en ratones BALB / c (Figura 9A) y en ratones C3B6 (Figura 9B) o luego de una dosis de cebado y una dosis de refuerzo de ARNm que codifica la proteína SARS-CoV-2 de tipo salvaje (Figura 9C). Se probaron varias dosis de vacuna. Las Figuras 9D-9E muestran gráficos que comparan los títulos de anticuerpos séricos en ratones BALB/c (Figura 9D) y en ratones C3B6 (Figura 9E) inmunizados con la vacuna de ARNm de la variante 9 del SARS-CoV-2, el ARNm de la vacuna de la variante 5 del SARS-CoV-2, o ARNm que codifica la proteína S del SARS-CoV-2 de tipo salvaje.

- La **Figura 10** muestra un gráfico que compara el título de anticuerpos en suero de ratones inmunizados con una de siete vacunas de ARNm de SARA-CoV-2 diferentes o un ARNm que codifica la secuencia de proteína S de SARS-CoV-2 de tipo salvaje luego de una dosis de refuerzo.

- Las **Figuras 11A-11B** muestran gráficos de los resultados de un análisis de citometría de flujo utilizando el anticuerpo 5653-118 ("118"), que es específico para el dominio N-terminal de la subunidad S1 de SARS-CoV-1, luego de la inmunización de ratones con la vacuna de ARNm de la variante 9 del SARS-CoV-2, la vacuna de ARNm de la variante 5 del SARS-CoV-2 o la vacuna de ARNm de la variante 6 del SARS-CoV-2. El análisis se realizó utilizando muestras de ganglios linfáticos (Figura 11A) y bazo (Figura 11B) obtenidas de los

ratones.

Las **Figuras 12A-12B** muestran gráficos de los resultados de un análisis de citometría de flujo utilizando el anticuerpo 5652-109 ("109"), que es específico para el dominio de unión a la proteína de S de SARS-CoV1, luego de la inmunización de ratones con la vacuna de ARNm de la variante 9 del SARS-CoV-2, la vacuna de ARNm de la variante 5 del SARS-CoV-2 o la vacuna de ARNm de la variante 6 del SARS-CoV-2. El análisis se realizó utilizando muestras de ganglio linfático (Figura 12A) y bazo (Figura 12B) obtenidas de los ratones.

Las **Figura 13A-13C** muestran gráficos de los resultados de un análisis de citometría de flujo luego de la transfección con una de las seis vacunas de ARNm de SARS-CoV-2 diferentes *in vitro*. La Figura 13A muestra el porcentaje de células presentadoras de antígeno positivas (APC +), y la Figura 13B muestra la intensidad de fluorescencia media (MFI). La Figura 13C muestra resultados utilizando un control positivo (un anticuerpo del SARS)

La **Figura 14** muestra gráficos de los resultados de un análisis de citometría de flujo luego de la transfección con la vacuna de ARNm de SARS-CoV-2 variante 9 *in vitro*, utilizando mAb118, mAb109 y mAb103 de SARS (control positivo). El control negativo excluyó un anticuerpo primario.

La **Figura 15** muestra gráficos de la unión de proteínas entre mAb118 o mAb109 y un antígeno del SARS-CoV-2 a diferentes concentraciones.

Las **Figuras 16A-16B** muestran gráficos de anticuerpos neutralizantes y de unión en ratones BALB/c vacunados con 1 µg, 0.1 µg o 0.01 µg de la vacuna de ARNm de la variante 9 del SARS-CoV-2 en las semanas 0 y 3. La Figura 16A muestra anticuerpos de unión a S-2P evaluados por ELISA en la semana 2 (postcebadado) y semana 5 (posrefuerzo). La Figura 16B muestra la actividad neutralizante evaluada en la semana 5 mediante un ensayo de neutralización de pseudovirus en sueros de ratones que recibieron 1 µg o 0.1 µg de la vacuna de ARNm de la variante 9 del SARS-CoV-2.

Las **Figuras 17A-17C** muestran gráficos de datos que demuestran que la inmunidad inducida por la vacuna del ARNm de la variante 9 del SARS-CoV-2 previene la replicación del SARS-CoV-2 en los pulmones de los ratones BALB/c. Se vacunaron ratones BALB/c con 1 µg, 0.1 µg o 0.01 µg de la vacuna de ARNm de la variante 9 del SARS-CoV-2 en las semanas 0 y 3 y se expusieron en la semana 9 con el SARS-CoV-2 adaptado al ratón. La

Figura 17A muestra títulos virales en pulmón evaluados mediante ensayo de placa el día 2 luego de la exposición. La Figura 17B muestra títulos virales en cornetes nasales evaluados por ensayo de placa el día 2 luego de la exposición. La Figura 17C muestra el cambio en el peso corporal (como porcentaje) a lo largo del tiempo luego de la infección.

- 5 Las **Figuras 18A-18C** muestran gráficos de datos que demuestran que la inmunidad inducida por la vacuna del ARNm de la variante 9 del SARS-CoV-2 previene la replicación del SARS-CoV-2 en los pulmones de los ratones BALB/c. Se vacunaron ratones BALB/c con 1 μg , 0.1 μg o 0.01 μg de la vacuna de ARNm de la variante 9 del SARS-CoV-2 en la semana 0 y se expusieron en la semana 7 con el SARS-CoV-2 adaptado al ratón. La Figura 10 18A muestra títulos virales en cornetes nasales evaluados por ensayo de placa el día 2 luego de la exposición. La Figura 18B muestra títulos virales en pulmón evaluados mediante ensayo de placa el día 2 luego de la exposición. La Figura 18C muestra el cambio en el peso corporal (como porcentaje) a lo largo del tiempo luego de la infección.

La **Figura 19** muestra el programa de inmunización de las semanas 0 y 3 utilizado en el 15 Ejemplo 10.

- Las **Figuras 20A-20C** muestran gráficos de datos que demuestran que la inmunidad inducida por la vacuna del ARNm de la variante 9 del SARS-CoV-2 previene la replicación del SARS-coV-2 en los pulmones de los ratones BALB/c. Se vacunaron ratones BALB/c con 10 μg , 1 μg o 0.1 μg de la variante 9 SARS-CoV-2 en las semanas 0 y 4 y en la semana 7 20 se expusieron a SARS-CoV-2 adaptado al ratón. La Figura 20A muestra títulos virales en cornetes nasales evaluados por ensayo de placa el día 2 luego de la exposición. La Figura 20B muestra títulos virales en pulmón evaluados mediante ensayo de placa el día 2 luego de la exposición. La Figura 20C muestra el cambio en el peso corporal (como porcentaje) a lo largo del tiempo luego de la infección.

- 25 Las **Figuras 21A-21H** muestran gráficos de datos relacionados con respuestas de anticuerpos neutralizantes luego de la inmunización de ARNm de ratones BALB/c. Se generaron curvas sigmoideas, tomando promedios de triplicados en cada dilución de suero, a partir de lecturas de unidades de luciferasa relativas (RLU) y se calculó la actividad neutralizante al 50 % (IC_{50}) (Figuras 21A, 21C, 21E, 21G) y 80 % (IC_{80}) (Figs. de 21B, 21D, 30 21F, 21H) considerando que las células no infectadas representan el 100 % de la neutralización y las células transducidas únicamente con virus para representar el 0 % de neutralización. Cada símbolo representa un ratón individual, las barras representan títulos

de media geométrica (GMT) y las barras de error indican la desviación estándar geométrica (SD). Las Figuras 21A-21F muestran pruebas T no emparejadas utilizadas para comparar dosis de 0.1 µg y 1 µg. Las Figuras 21G y 21H muestran grupos comparados mediante ANOVA unidireccional con la prueba de comparación múltiple de Kruskal-Wallis.

- 5 Las **Figuras 22A-22C** muestran gráficos de datos relacionados con la unión y las respuestas de anticuerpos neutralizantes luego de la inmunización de ARNm de dosis baja de ratones BALB/c con diseños de antígenos de punta alternativos. La Figura 22A muestra títulos de punto final en suero. La Figura 22B muestra el 50 % (IC₅₀) de actividad neutralizante calculada considerando las células no infectadas que representan 100 % de neutralización y las células transducidas únicamente con el virus que representa 0 % de neutralización. Cada símbolo representa un ratón individual, las barras representan títulos de media geométrica (GMT) y las barras de error indican la desviación estándar geométrica (SD). En las Figuras 22A y 22B, los grupos se compararon mediante ANOVA de una vía con la prueba de comparación múltiple de Kruskal-Wallis. La Figura 22C muestra títulos de neutralización y unión de anticuerpos comparados por correlación de Spearman.

DESCRIPCIÓN DETALLADA

- La presente descripción proporciona composiciones (p. ej., composiciones inmunizantes/inmunogénicas tales como vacunas de ARN en nanopartículas lipídicas) que provocan potentes anticuerpos neutralizantes contra antígenos de coronavirus. En algunas modalidades, una composición inmunizante incluye ARN (p. ej., ARN mensajero (ARNm)) que codifica un antígeno de coronavirus, como un antígeno de SARS-CoV-2 en una nanopartícula lipídica. En algunas modalidades, el antígeno de coronavirus es una proteína estructural. En algunas modalidades, el antígeno de coronavirus es una proteína de pico, una proteína de envoltura, una proteína de nucleocápside o una proteína de membrana. En algunas modalidades, el antígeno de coronavirus es una proteína de pico de prefusión estabilizada. En algunas modalidades, el ARNm comprende un marco de lectura abierto que codifica una proteína de pico trimérica variante. Por ejemplo, la proteína del pico trimérica puede comprender una proteína del pico de prefusión estabilizada. En algunas modalidades, la proteína de pico de prefusión estabilizada tiene una mutación de prolina doble (S2P).

Antígenos

Los antígenos, como se utilizan en este documento, son proteínas capaces de inducir una respuesta inmune (p. ej., hacer que un sistema inmune produzca anticuerpos contra los antígenos). En este documento, el uso del término "antígeno" abarca proteínas inmunogénicas y fragmentos inmunogénicos (un fragmento inmunogénico que induce (o es capaz de inducir) una respuesta inmune a un (al menos un) coronavirus), a menos que se indique lo contrario. Debe entenderse que el término "proteína" abarca péptidos y el término "antígeno" abarca fragmentos antigénicos. Otras moléculas pueden ser antigénicas tales como polisacáridos bacterianos o combinaciones de estructuras de proteínas y polisacáridos, pero para las vacunas virales incluidas en este documento, las proteínas virales, fragmentos de proteínas virales y proteínas diseñadas o mutadas derivadas del betacoronavirus SARS-CoV-2 son los antígenos proporcionados en este documento.

El ARNm proporcionado en la presente, en algunas modalidades, comprende un marco de lectura abierto que codifica una variante de proteína de pico trimérica. En algunas modalidades, el marco de lectura abierto codifica una variante de proteína de pico trimérica que comprende una proteína de pico de prefusión estabilizada. La proteína de pico de prefusión estabilizada, en algunas modalidades, comprende una mutación de prolina doble (S2P).

Las secuencias ilustrativas de los antígenos de coronavirus y el ARN (p. ej., ARNm) que codifican los antígenos de coronavirus de las composiciones de la presente descripción se proporcionan en la Tabla 1.

En algunas modalidades, una composición comprende un ARN (p. ej., ARNm) que codifica un antígeno de coronavirus que comprende una secuencia de aminoácidos que tiene al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 98 %, o 100 % de identidad con la secuencia de aminoácidos de cualquiera de las SEQ ID NO: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83 u 85 y opcionalmente una nanopartícula lipídica. En algunas modalidades, una composición comprende un ARN (p. ej., ARNm) que codifica un antígeno de coronavirus que comprende la secuencia de SEQ ID NO: 29. En algunas modalidades, una composición comprende un ARN (p. ej., ARNm) que codifica un antígeno de coronavirus que comprende la secuencia de SEQ ID NO: 17. En algunas modalidades, una composición comprende un ARN (p. ej., ARNm) que codifica un antígeno de coronavirus que comprende la secuencia de SEQ ID NO: 20. En algunas modalidades, una composición comprende un ARN (p. ej., ARNm) que

codifica un antígeno de coronavirus que comprende la secuencia de SEQ ID NO: 23. En algunas modalidades, una composición comprende un ARN (p. ej., ARNm) que codifica un antígeno de coronavirus que comprende la secuencia de SEQ ID NO: 26.

5 Debe entenderse que cualquiera de los antígenos codificados por el ARN descrito en el presente documento puede comprender o no una secuencia señal.

Ácidos nucleicos

10 Las composiciones de la presente descripción comprenden un (al menos un) ARN que tiene un marco de lectura abierto (ORF) que codifica un antígeno de coronavirus (p. ej., proteína de pico trimérica variante, como una proteína de pico de prefusión estabilizada). En algunas modalidades, el ARN es un ARN mensajero (ARNm). En algunas modalidades, el ARN (p. ej., ARNm) comprende además una UTR 5', UTR 3', una cola poli(A) y/o un análogo de cubierta 5'.

15 También debe entenderse que la vacuna contra el coronavirus de la presente descripción puede incluir cualquier región 5' no traducida (UTR) y/o cualquier 3' UTR. Se proporcionan secuencias de UTR ilustrativas en el Listado de Secuencias (p. ej., SEQ ID NO: 2, 36, 4 o 37); sin embargo, se pueden utilizar o intercambiar otras secuencias de UTR por cualquiera de las secuencias de UTR descritas en este documento. Las UTR también pueden omitirse de los polinucleótidos de ARN proporcionados en este documento.

20 Los ácidos nucleicos comprenden un polímero de nucleótidos (monómeros de nucleótidos). Por lo tanto, los ácidos nucleicos también se denominan polinucleótidos. Los ácidos nucleicos pueden ser o pueden incluir, por ejemplo, ácidos desoxirribonucleicos (ADN), ácidos ribonucleicos (ARN), ácidos nucleicos treosa (TNA), ácidos nucleicos de glicol (GNA), ácidos nucleicos peptídicos (PNA), ácidos nucleicos bloqueados (LNA), incluyendo 25 LNA que tiene una configuración β -D-ribo, α -LNA que tiene una configuración α -L-ribo (un diastereómero de LNA), 2'-amino-LNA que tiene una funcionalización 2'-amino y 2'-amino- α -LNA que tiene una funcionalización 2'-amino), ácidos etilenucleicos (ENA), ácidos ciclohexenil nucleicos (CeNA) y/o quimeras y/o combinaciones de estos.

30 El ARN mensajero (ARNm) es cualquier ARN que codifica una (al menos una) proteína (un polímero de aminoácidos de origen natural, no natural o modificado) y puede traducirse para producir la proteína codificada *in vitro*, *in vivo*, *in situ* o *ex vivo*. El experto en la técnica apreciará que, salvo que se indique lo contrario, las secuencias de ácido nucleico

establecidas en la presente solicitud pueden recitar "T" en una secuencia de ADN representativa, pero cuando la secuencia representa ARN (p. ej., ARNm), las "T" se sustituirán por "U". Por lo tanto, cualquiera de los ADN descritos e identificados por un número de identificación de secuencia particular en este documento también describe la

5 secuencia de ARN correspondiente (p. ej., ARNm) complementaria al ADN, donde cada "T" de la secuencia de ADN está sustituida con "U".

Un marco de lectura abierto (ORF) es un tramo continuo de ADN o ARN que comienza con un codón de inicio (p. ej., metionina (ATG o AUG)) y termina con un codón de terminación (p. ej., TAA, TAG o TGA, o UAA, UAG o UGA). Un ORF normalmente codifica una proteína.

10 Se entenderá que las secuencias descritas en el presente documento pueden comprender además elementos adicionales, p. ej., 5' y 3' UTR, pero que esos elementos, a diferencia del ORF, no necesitan estar presentes necesariamente en un polinucleótido de ARN de la presente descripción.

En algunas modalidades, una composición comprende un ARN (p. ej., ARNm) que

15 comprende una secuencia de nucleótidos que tiene al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 98 % o 100 % de identidad a la secuencia de nucleótidos de cualquiera de las SEQ ID NO: 1, 6, 9, 12, 15, 18, 21, 24, 27, 30, 51, 53, 55, 57, 58, 60 u 86-97.

En algunas modalidades, una composición comprende un ARN (p. ej., ARNm) que

20 comprende un ORF que comprende una secuencia de nucleótidos que tiene al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 98 %, o 100 % de identidad con la secuencia de nucleótidos de cualquiera de las SEQ ID NO: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 u 84.

25 **Variantes**

En algunas modalidades, las composiciones de la presente descripción incluyen ARN que codifica una variante de antígeno de coronavirus (p. ej., proteína de pico trimérica variante, como una proteína de pico de prefusión estabilizada). Las variantes de antígeno u otras

variantes de polipéptidos se refieren a moléculas que difieren en su secuencia de

30 aminoácidos de una secuencia de referencia, nativa o de tipo salvaje. Las variantes de antígeno/polipéptido pueden poseer sustituciones, eliminaciones y/o inserciones en ciertas posiciones dentro de la secuencia de aminoácidos, en comparación con una secuencia

nativa o de referencia. Normalmente, las variantes poseen al menos un 50 % de identidad con una secuencia de referencia, nativa o de tipo salvaje. En algunas modalidades, las variantes comparten al menos el 80 % o al menos el 90 % de identidad con una secuencia de referencia, nativa o de tipo salvaje.

- 5 Los antígenos/polipéptidos variantes codificados por los ácidos nucleicos de la descripción pueden contener cambios de aminoácidos que confieren una serie de propiedades deseables, p. ej., que mejoran su inmunogenicidad, mejoran su expresión y/o mejoran su estabilidad o propiedades PK/PD [farmacocinética(s)/farmacodinámica(s)] en un sujeto. Pueden prepararse antígenos/polipéptidos variantes utilizando técnicas de mutagénesis de rutina y ensayarse según sea apropiado para determinar si poseen la propiedad deseada. Los ensayos para determinar los niveles de expresión y la inmunogenicidad son bien conocidos en la técnica y en la sección de Ejemplos se exponen ejemplos de tales ensayos. De manera similar, las propiedades PK/PD de una variante de proteína se pueden medir utilizando técnicas reconocidas en la técnica, p. ej., determinando la expresión de antígenos en un sujeto vacunado a lo largo del tiempo y/u observando la durabilidad de la respuesta inmune inducida. La estabilidad de la o las proteínas codificadas por una variante de ácido nucleico puede medirse ensayando la estabilidad térmica o la estabilidad tras la desnaturalización de la urea o puede medirse utilizando predicción in silico. Los métodos para tales experimentos y determinaciones in silico son conocidos en la técnica.
- 10 En algunas modalidades, una composición comprende un ARN o un ORF de ARN que comprende una secuencia de nucleótidos de cualquiera de las secuencias proporcionadas en este documento (véase, p. ej., Listado de secuencias y Tabla 1), o comprende una secuencia de nucleótidos al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % idéntica a una secuencia de nucleótidos de cualquiera de las secuencias proporcionadas en este documento.
- 15 El término "identidad" se refiere a una relación entre las secuencias de dos o más polipéptidos (p. ej., antígenos) o polinucleótidos (ácidos nucleicos), según se determina comparando las secuencias. La identidad también se refiere al grado de relación de secuencia entre secuencias determinado por el número de coincidencias entre cadenas de dos o más residuos de aminoácidos o residuos de ácido nucleico. La identidad mide el porcentaje de coincidencias idénticas entre la menor de dos o más secuencias con alineaciones de espacios (si las hay) abordadas por un modelo matemático o programa
- 20
- 25
- 30

informático en particular (p. ej., "algoritmos"). La identidad de los antígenos o ácidos nucleicos relacionados se puede calcular fácilmente mediante métodos conocidos. El "porcentaje (%) de identidad de secuencia de aminoácidos" con respecto a una secuencia de polipéptidos de referencia se define como el porcentaje de residuos de aminoácidos en una secuencia candidata que son idénticos a los residuos de aminoácidos en la secuencia de polipéptidos de referencia, luego de alinear las secuencias e introducir espacios, si fuera necesario, para lograr el porcentaje máximo de identidad de secuencia y sin considerar cualquier sustitución conservadora como parte de la identidad de secuencia. Los métodos y programas informáticos para la alineación son bien conocidos en la técnica. Se entiende que la identidad depende de un cálculo del porcentaje de identidad, pero puede diferir en valor debido a carencias y penalizaciones introducidas en el cálculo. Generalmente, las variantes de un polinucleótido o polipéptido particular (p. ej., Antígeno) tienen al menos 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % pero menos del 100 % de identidad de secuencia con ese polinucleótido o polipéptido de referencia particular según lo determinado por los programas y parámetros de alineación de secuencia descritos en este documento y conocido por los expertos en la técnica. Tales herramientas para la alineación incluyen las del paquete BLAST (Stephen F. Altschul, et al (1997), "Gapped BLAST y PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402). Otra técnica de alineación local popular se basa en el algoritmo de Smith-Waterman (Smith, TF & Waterman, MS (1981) "Identification of common molecular subsequences." *J. Mol. Biol.* 147: 195-197). Una técnica de alineación global general basada en la programación dinámica es el algoritmo de Needleman-Wunsch (Needleman, SB & Wunsch, CD (1970) "A general method applicable to the search for similarities in the amino acid sequences of two proteins." *J. Mol. Biol.* 48: 443-453). Más recientemente, se ha desarrollado un algoritmo de alineación de secuencia global óptima rápida (FOGSAA) que supuestamente produce una alineación global de secuencias de nucleótidos y proteínas más rápido que otros métodos de alineación global óptima, incluido el algoritmo Needleman-Wunsch.

Como tales, los polinucleótidos que codifican péptidos o polipéptidos que contienen sustituciones, inserciones y/o adiciones, eliminaciones y modificaciones covalentes con respecto a las secuencias de referencia, en particular las secuencias de polipéptidos (p. ej., antígenos) descritas en este documento, se incluyen dentro del alcance de esta descripción.

Por ejemplo, se pueden añadir etiquetas de secuencia o aminoácidos, tales como una o más lisinas, a secuencias de péptidos (p. ej., en los extremos N-terminal o C-terminal). Las etiquetas de secuencia se pueden utilizar para la detección, purificación o localización de péptidos. Se pueden utilizar lisinas para aumentar la solubilidad de los péptidos o para permitir la biotinilación. Alternativamente, los residuos de aminoácidos ubicados en las regiones carboxi y amino terminales de la secuencia de aminoácidos de un péptido o proteína se pueden eliminar opcionalmente proporcionando secuencias truncadas. Determinados aminoácidos (p. ej., residuos C-terminales o N-terminales) se pueden eliminar alternativamente dependiendo del uso de la secuencia, como por ejemplo, la expresión de la secuencia como parte de una secuencia más grande que es soluble o está unida a un soporte sólido. En algunas modalidades, las secuencias para (o que codifican) las secuencias señal, secuencias de terminación, dominios transmembrana, enlazadores, dominios de multimerización (tales como, p. ej., regiones de pliegue) y similares pueden sustituirse con secuencias alternativas que logran la misma función o una función similar. En algunas modalidades, las cavidades en el núcleo de las proteínas pueden rellenarse para mejorar la estabilidad, p. ej., introduciendo aminoácidos más grandes. En otras modalidades, las redes de enlaces de hidrógeno enterradas se pueden reemplazar con residuas hidrófobas para mejorar la estabilidad. En otras modalidades adicionales, los sitios de glicosilación pueden eliminarse y reemplazarse con residuos apropiados. Dichas secuencias son fácilmente identificables por un experto en la técnica. También debe entenderse que algunas de las secuencias proporcionadas en este documento contienen etiquetas de secuencia o secuencias de péptidos terminales (p. ej., en los extremos N-terminal o C-terminal) que se pueden eliminar, por ejemplo, antes de su uso en la preparación de una vacuna de ARN (p. ej., ARNm). Como reconocen los expertos en la técnica, los fragmentos de proteínas, los dominios de proteínas funcionales y las proteínas homólogas también se consideran dentro del alcance de los antígenos de coronavirus de interés. Por ejemplo, aquí se proporciona cualquier fragmento de proteína (es decir, una secuencia polipeptídica de al menos un residuo de aminoácido más corta que una secuencia de antígeno de referencia, pero por lo demás idéntica) de una proteína de referencia, siempre que el fragmento sea inmunogénico y confiera una respuesta inmune protectora al coronavirus. Además de las variantes que son idénticas a la proteína de referencia pero están truncadas, en algunas modalidades, un

antígeno incluye 2, 3, 4, 5, 6, 7, 8, 9, 10 o más mutaciones, como se muestra en cualquiera de las secuencias proporcionadas o referenciadas en este documento. Los antígenos/polipéptidos antigénicos pueden variar en longitud desde aproximadamente 4, 6 u 8 aminoácidos hasta proteínas de longitud completa.

5

Elementos estabilizadores

Las moléculas de ARNm eucariotas de origen natural pueden contener elementos estabilizadores, que incluyen, entre otros, regiones no traducidas (UTR) en su extremo 5' (5' UTR) y/o en su extremo 3' (3' UTR), además a otras características estructurales, como una estructura de cubierta 5' o una cola de 3'-poli(A). Tanto la 5' UTR como la 3' UTR se transcriben típicamente del ADN genómico y son elementos del ARNm prematuro. Los rasgos estructurales característicos del ARNm maduro, como la cubierta 5' y la cola 3'-poli(A), se añaden habitualmente al ARNm transcrito (prematuro) durante el procesamiento del ARNm.

En algunas modalidades, una composición incluye un polinucleótido de ARN que tiene un marco de lectura abierto que codifica al menos un polipéptido antigénico que tiene al menos una modificación, al menos una tapa terminal 5', y se formula dentro de una nanopartícula lipídica. La cubierta 5' de polinucleótidos se puede completar de forma concomitante durante la reacción de transcripción *in vitro* utilizando los siguientes análogos químicos de cubierta de ARN para generar la estructura de la cubierta 5'-guanosina de conformidad con los protocolos del fabricante: 3'-O-Me-m7G(5')ppp(5') G [the ARCA cap]; G(5')ppp(5')A; G(5')ppp(5')G; m7G(5')ppp(5')A; m7G(5')ppp(5')G (New England BioLabs, Ipswich, MA). La cubierta 5' del ARN modificado puede completarse postranscripcionalmente utilizando una enzima de cubierta del virus Vaccinia para generar la estructura "Cap 0": m7G(5')ppp(5')G (New England BioLabs, Ipswich, MA). La estructura de Cap 1 se puede generar utilizando tanto la Enzima de Cubierta del Virus Vaccinia como una 2'-O metil-transferasa para generar: m7G(5')ppp(5')G-2'-O-metilo. La estructura Cap 2 puede generarse a partir de la estructura Cap 1 seguida de la metilación 2'-O del antepenúltimo nucleótido 5' utilizando una 2'-O metil-transferasa. La estructura de Cap 3 puede generarse a partir de la estructura de Cap 2 seguida de la metilación 2'-O del antepenúltimo nucleótido 5' utilizando una 2'-O metil-transferasa. Las enzimas pueden derivarse de una fuente recombinante.

La cola 3'-poli(A) es típicamente un tramo de nucleótidos de adenina añadidos al extremo

3' del ARNm transcrito. En algunos casos, puede comprender hasta aproximadamente 400 nucleótidos de adenina. En algunas modalidades, la longitud de la cola 3'-poli(A) puede ser un elemento esencial con respecto a la estabilidad del ARNm individual.

En algunas modalidades, una composición incluye un elemento estabilizador. Los elementos estabilizadores pueden incluir, por ejemplo, un tallo-bucle de histona. Se ha identificado una proteína de unión de tallo-bucle (SLBP), una proteína de 32 kDa. Está asociado con el tallo-bucle de histonas en el extremo 3' de los mensajes de histonas tanto en el núcleo como en el citoplasma. Su nivel de expresión está regulado por el ciclo celular; alcanza su punto máximo durante la fase S, cuando los niveles de ARNm de histonas también están elevados. Se ha demostrado que la proteína es esencial para el procesamiento eficaz en el extremo 3' del pre-ARNm de histona por parte del snRNP de U7. La SLBP sigue estando asociada con el tallo-bucle luego del procesamiento y luego estimula la traducción de los ARNm de histonas maduras en proteínas de histonas en el citoplasma. El dominio de unión de ARN de SLBP se conserva a través de metazoos y protozoos; su unión al tallo-bucle de histona depende de la estructura del bucle. El sitio de unión mínimo incluye al menos tres nucleótidos 5' y dos nucleótidos 3' con respecto al tallo-bucle.

En algunas modalidades, un ARN (p. ej., ARNm) incluye una región codificante, al menos un tallo-bucle de histona y, opcionalmente, una secuencia de poli(A) o señal de poliadenilación. La secuencia de poli(A) o señal de poliadenilación generalmente debería mejorar el nivel de expresión de la proteína codificada. La proteína codificada, en algunas modalidades, no es una proteína histona, una proteína indicadora (p. ej. luciferasa, GFP, EGFP, β -galactosidasa, EGFP), o un marcador o proteína de selección (p. ej., alfa-globina, galactoquinasa y xantina: guanina fosforribosil transferasa (GPT)).

En algunas modalidades, un ARN (p. ej., ARNm) incluye la combinación de una secuencia de poli(A) o señal de poliadenilación y al menos un tallo-bucle de histona, si bien ambos representan mecanismos alternativos en la naturaleza, actúan sinérgicamente para aumentar la expresión de la proteína más allá el nivel observado con cualquiera de los elementos individuales. El efecto sinérgico de la combinación de poli(A) y al menos un tallo-bucle de histona no depende del orden de los elementos o de la longitud de la secuencia de poli(A).

En algunas modalidades, un ARN (p. ej., ARNm) no incluye un elemento de histona corriente abajo (HDE). El "elemento de histona corriente abajo" (HDE) incluye un tramo de

polinucleótidos rico en purina de aproximadamente 15 a 20 nucleótidos 3' de tallo-lazos naturales, que representan el sitio de unión para el ARNnp (snARN) U7, que está involucrado en el procesamiento de preARNm de histonas en ARNm de histona madura. En algunas modalidades, el ácido nucleico no incluye un intrón.

- 5 Un ARN (p. ej., ARNm) puede contener o no una secuencia potenciadora y/o promotora, que puede estar modificada o no modificada o que puede estar activada o inactivada. En algunas modalidades, el tallo-bucle de histonas se deriva generalmente de genes de histonas e incluye un emparejamiento de bases intramoleculares de dos secuencias complementarias inversas vecinas parcial o totalmente separadas por un espaciador, que
10 consta de una secuencia corta, que forma el bucle de la estructura. La región de bucle no emparejado normalmente no puede emparejarse con ninguno de los elementos de tallo-bucle. Ocurre con más frecuencia en el ARN, ya que es un componente clave de muchas estructuras secundarias de ARN, pero también puede estar presente en el ADN monocatenario. La estabilidad de la estructura del tallo-bucle generalmente depende de la
15 longitud, el número de desajustes o protuberancias y la composición de la base de la región emparejada. En algunas modalidades, puede resultar un emparejamiento de bases oscilantes (emparejamiento de bases que no son de Watson-Crick). En algunas modalidades, la al menos una secuencia de tallo-bucle de histona comprende una longitud de 15 a 45 nucleótidos.
- 20 En algunas modalidades, un ARN (p. ej., ARNm) tiene una o más secuencias ricas en AU eliminadas. Estas secuencias, a veces denominadas AURES, son secuencias desestabilizadoras que se encuentran en la 3' UTR. Las AURES pueden eliminarse de las vacunas de ARN. Alternativamente, las AURES pueden permanecer en la vacuna de ARN.

25 **Péptidos de señal**

- En algunas modalidades, una composición comprende un ARN (p. ej., ARNm) que tiene un ORF que codifica un péptido señal fusionado con el antígeno del coronavirus. Los péptidos señal, que comprenden los 15-60 aminoácidos N-terminales de las proteínas, son típicamente necesarios para la translocación a través de la membrana en la vía secretora y,
30 por lo tanto, controlan universalmente la entrada de la mayoría de las proteínas en eucariotas y procariotas a la vía secretora. En eucariotas, el péptido señal de una proteína precursora naciente (preproteína) dirige el ribosoma a la membrana del retículo

endoplásmico rugoso (RE) e inicia el transporte de la cadena de péptidos en crecimiento a través de esta para su procesamiento. El procesamiento de ER produce proteínas maduras, en las que el péptido señal se escinde de las proteínas precursoras, típicamente por una peptidasa señal residente en ER de la célula huésped, o permanecen sin escindir y funcionan como un ancla de membrana. Un péptido señal también puede facilitar el direccionamiento de la proteína a la membrana celular.

Un péptido señal puede tener una longitud de 15 a 60 aminoácidos. Por ejemplo, un péptido señal puede tener una longitud de 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, o 60 aminoácidos. En algunas modalidades, un péptido señal tiene una longitud de 20-60, 25-60, 30-60, 35-60, 40-60, 45-60, 50-60, 55-60, 15-55, 20-55, 25-55, 30-55, 35-55, 40-55, 45-55, 50-55, 15-50, 20-50, 25-50, 30-50, 35-50, 40-50, 45-50, 15-45, 20-45, 25-45, 30-45, 35-45, 40-45, 15-40, 20-40, 25-40, 30-40, 35-40, 15-35, 20-35, 25-35, 30-35, 15-30, 20-30, 25-30, 15-25, 20-25 o 15-20 aminoácidos.

Los péptidos señal de genes heterólogos (que regulan la expresión de genes distintos de los antígenos de coronavirus en la naturaleza) son conocidos en la técnica y pueden probarse para determinar las propiedades deseadas y luego incorporarse en un ácido nucleico de la descripción. En algunas modalidades, el péptido señal puede comprender una de las siguientes secuencias:

MDSKGSSQKGSRLLLLLVSNLLLPQGVVG (SEQ ID NO: 38), MDWTWILFLVAAATRVHS (SEQ ID NO: 39); METPAQLLFLLLLWLPDTTG (SEQ ID NO: 40); MLGSNSGQRVVFTILLLLVAPAYS (SEQ ID NO: 41); MKCLLYLAFLFIGVNCA (SEQ ID NO: 42); MWLVSLAIVTACAGA (SEQ ID NO: 43); o MFVFLVLLPLVSSQC (SEQ ID NO: 99).

Proteínas de fusión

En algunas modalidades, una composición de la presente descripción incluye un ARN (p. ej., ARNm) que codifica una proteína de fusión antigénica. Por lo tanto, el antígeno o los antígenos codificados pueden incluir dos o más proteínas (p. ej., proteína y/o fragmento de proteína) unidas. Alternativamente, la proteína a la que se fusiona un antígeno proteico no promueve una fuerte respuesta inmune a sí misma, sino al antígeno del coronavirus. Las proteínas de fusión antigénicas, en algunas modalidades, retienen la propiedad funcional de cada proteína original.

Restos de armazones proteicos

Las vacunas de ARN (p. ej., ARNm) como se proporciona en el presente documento, en algunas modalidades, codifican proteínas de fusión que comprenden antígenos de coronavirus unidos a restos de armazones proteicos. En algunas modalidades, tales restos de armazón imparten las propiedades deseadas a un antígeno codificado por un ácido nucleico de la descripción. Por ejemplo, las proteínas de armazón pueden mejorar la inmunogenicidad de un antígeno, p. ej., alterando la estructura del antígeno, alterando la captación y procesamiento del antígeno y/o haciendo que el antígeno se una a un compañero de unión.

En algunas modalidades, el resto de armazones proteicos es una proteína que puede autoensamblarse en nanopartículas de proteínas que son altamente simétricas, estables y están organizadas estructuralmente, con diámetros de 10 a 150 nm, un rango de tamaño muy adecuado para interacciones óptimas con varias células del sistema inmunitario. En algunas modalidades, pueden utilizarse proteínas virales o partículas similares a virus para formar estructuras de nanopartículas estables. En la técnica se conocen ejemplos de tales proteínas virales. Por ejemplo, en algunas modalidades, el resto de armazones proteicos es un antígeno de superficie de la hepatitis B (HBsAg). El HBsAg forma partículas esféricas con un diámetro promedio de ~22 nm y que carecían de ácido nucleico y, por lo tanto, no son infecciosas (Lopez-Sagaseta, J. et al. Computational and Structural Biotechnology Journal 14 (2016) 58–68). En algunas modalidades, el resto de armazón proteico es un antígeno del núcleo de la hepatitis B (HBcAg) que se autoensambla en partículas de 24-31 nm de diámetro, que se asemejan a los núcleos virales obtenidos de hígado humano infectado por HBV. El HBcAg producido por autoensamblaje en dos clases de nanopartículas de diferentes tamaños de 300 Å y 360 Å de diámetro, correspondientes a 180 o 240 protómeros. En algunas modalidades, el antígeno de coronavirus se fusiona con HBsAG o HBcAG para facilitar el autoensamblaje de nanopartículas que presentan el antígeno de coronavirus.

En algunas modalidades, se pueden utilizar plataformas de proteínas bacterianas. Los ejemplos no taxativos de estas proteínas de autoensamblaje incluyen ferritina, lumazina y encapsulina.

La ferritina es una proteína cuya función principal es el almacenamiento de hierro

intracelular. La ferritina está compuesta por 24 subunidades, cada una compuesta por un haz de cuatro hélices alfa, que se autoensamblan en una estructura cuaternaria con simetría octaédrica (Cho KJ. *Et al. J Mol Biol.* 2009; 390:83–98). Se han determinado varias estructuras de ferritina de alta resolución, lo que confirma que la ferritina de *Helicobacter pylori* está compuesta por 24 protómeros idénticos, mientras que en los animales hay cadenas ligeras y pesadas de ferritina que pueden ensamblarse solas o combinarse con diferentes proporciones en partículas de 24 subunidades T. *y col., J Biol Inorg Chem.* 2003;8:105-111; Lawson DM *et al. Nature.* 1991;349:541-544). La ferritina se autoensambla en nanopartículas con una sólida estabilidad térmica y química. Por lo tanto, la nanopartícula de ferritina es adecuada para transportar y exponer antígenos.

La lumazina sintasa (LS) también es adecuada como plataforma de nanopartículas para la visualización de antígenos. LS, que es responsable del penúltimo paso catalítico en la biosíntesis de riboflavina, es una enzima presente en una amplia variedad de organismos, que incluyen arqueas, bacterias, hongos, plantas y eubacterias (Weber SE Flavins y Flavoproteins. *Methods and Protocols, Serie: Methods in Molecular Biology.* 2014). El monómero LS tiene 150 aminoácidos de longitud y consta de hojas beta junto con hélices alfa en tándem que flanquean sus lados. Se han informado varias estructuras cuaternarias diferentes para LS, lo que ilustra su versatilidad morfológica: desde homopentámeros hasta ensamblajes simétricos de 12 pentámeros que forman cápsides de 150 Å de diámetro. Incluso se han descrito jaulas LS de más de 100 subunidades (Zhang X. *et al. J Mol Biol.* 2006; 362:753–770).

La encapsulina, una nueva nanopartícula de jaula de proteínas aislada del termófilo *Thermotoga maritima*, también se puede utilizar como plataforma para presentar antígenos en la superficie de nanopartículas autoensambladas. La encapsulina se ensambla a partir de 60 copias de monómeros idénticos de 31 kDa que tienen una estructura de jaula simétrica delgada e icosaédrica T = 1 con diámetros interior y exterior de 20 y 24 nm, respectivamente (Sutter M. *et al. Nat Struct Mol Biol.* 2008, 15:939-947). Aunque la función exacta de la encapsulina en *T. maritima* aún no se comprende claramente, su estructura cristalina se ha resuelto recientemente y su función se postuló como un compartimento celular que encapsula proteínas como DyP (peroxidasa decolorante del tinte) y Flp (proteína similar a la ferritina), que están implicados en las respuestas al estrés oxidativo (Rahmanpour R. *et al. FEBS J.* 2013, 280: 2097-2104).

En algunas modalidades, un ARN de la presente descripción codifica un antígeno de coronavirus (p. ej., proteína SARS-CoV-2 S) fusionado a un dominio de pliegue. El dominio de pliegue puede obtenerse, por ejemplo, de fibritina del bacteriófago T4 (véase, p. ej., Tao Y, *et al. Structure*. 15 de junio de 1997; 5(6):789-98).

5

Enlazadores y péptidos escindibles

En algunas modalidades, los ARNm de la descripción codifican más de un polipéptido, denominado en el presente documento como proteínas de fusión. En algunas modalidades, el ARNm codifica además un enlazador ubicado entre al menos uno o cada dominio de la proteína de fusión. El enlazador puede ser, por ejemplo, un enlazador escindible o un enlazador sensible a proteasas. En algunas modalidades, el enlazador se selecciona del grupo que consiste en enlazador F2A, enlazador P2A, enlazador T2A, enlazador E2A y combinaciones de estos. Esta familia de enlazadores de péptidos autodescindibles, denominados péptidos 2A, se ha descrito en la técnica (véase, por ejemplo, Kim, JH et al. (2011) *PLoS ONE* 6:e18556). En algunas modalidades, el enlazador es un enlazador F2A. En algunas modalidades, el enlazador es un enlazador GGGS (SEQ ID NO: 98). En algunas modalidades, la proteína de fusión contiene tres dominios con enlazadores que intervienen, que tienen la estructura: dominio-enlazador-dominio-enlazador-dominio.

Se pueden utilizar enlazadores escindibles conocidos en la técnica en relación con la descripción. Los ejemplos de tales enlazadores incluyen: enlazadores F2A, enlazadores T2A, enlazadores P2A, enlazadores E2A (véase, p. ej., WO2017/127750). El experto en la técnica apreciará que otros enlazadores reconocidos en la técnica pueden ser adecuados para su uso en las construcciones de la descripción (p. ej., codificados por los ácidos nucleicos de la descripción). El experto en la técnica apreciará igualmente que otras construcciones policistrónicas (ARNm que codifica más de un antígeno/polipéptido por separado dentro de la misma molécula) pueden ser adecuadas para su uso, como se proporciona en el presente documento.

Optimización de la secuencia

En algunas modalidades, un ORF que codifica un antígeno de la descripción tiene un codón optimizado. Los métodos de optimización de codones son conocidos en la técnica. Por ejemplo, un ORF de cualquiera o más de las secuencias proporcionadas en este documento

puede tener codones optimizados. La optimización de codones, en algunas modalidades, puede utilizarse para hacer coincidir las frecuencias de codones en organismos diana y hospedador para asegurar un plegamiento apropiado; sesgar el contenido de GC para aumentar la estabilidad del ARNm o reducir las estructuras secundarias; minimizar los codones de repetición en tándem o las series de bases que pueden perjudicar la construcción o expresión de genes; personalizar las regiones de control de la transcripción y la traducción; insertar o eliminar secuencias de tráfico de proteínas; eliminar/añadir sitios de modificación posteriores a la traducción en la proteína codificada (p. ej., sitios de glicosilación); añadir, eliminar o mezclar dominios proteicos; insertar o eliminar sitios de restricción; modificar los sitios de unión del ribosoma y los sitios de degradación del ARNm; ajustar las velocidades de traducción para permitir que los diversos dominios de la proteína se plieguen correctamente; o reducir o eliminar las estructuras secundarias problemáticas dentro del polinucleótido. Las herramientas, algoritmos y servicios de optimización de codones son conocidos en la técnica; los ejemplos no taxativos incluyen servicios de GeneArt (Life Technologies), DNA2.0 (Menlo Park CA) y/o métodos patentados. En algunas modalidades, la secuencia del marco de lectura abierto (ORF) se optimiza utilizando algoritmos de optimización.

En algunas modalidades, una secuencia de codón optimizado comparte menos del 95 % de identidad de secuencia con un ORF de secuencia natural o de tipo salvaje (p. ej., una secuencia de ARNm de origen natural o de tipo salvaje que codifica un antígeno de coronavirus). En algunas modalidades, una secuencia de codones optimizados comparte menos del 90 % de identidad de secuencia con una secuencia de origen natural o de tipo salvaje (p. ej., una secuencia de ARNm de origen natural o de tipo salvaje que codifica un antígeno de coronavirus). En algunas modalidades, una secuencia de codón optimizado comparte menos del 85 % de identidad de secuencia con una secuencia de origen natural o de tipo salvaje (p. ej., una secuencia de ARNm de origen natural o de tipo salvaje que codifica un antígeno de coronavirus). En algunas modalidades, una secuencia de codón optimizado comparte menos del 80 % de identidad de secuencia con una secuencia de origen natural o de tipo salvaje (p. ej., una secuencia de ARNm de origen natural o de tipo salvaje que codifica un antígeno de coronavirus). En algunas modalidades, una secuencia de codones optimizados comparte menos del 75 % de identidad de secuencia con una secuencia de origen natural o de tipo salvaje (p. ej., una secuencia de ARNm de origen

natural o de tipo salvaje que codifica un antígeno de coronavirus).

En algunas modalidades, una secuencia de codones optimizados comparte entre 65 % y 85 % (p. ej., entre aproximadamente 67 % y aproximadamente 85 % o entre aproximadamente 67 % y aproximadamente 80 %) de identidad de secuencia con una secuencia de origen natural o de tipo salvaje (p. ej., una secuencia de ARNm de origen natural o de tipo salvaje que codifica un antígeno de coronavirus). En algunas modalidades, una secuencia de codones optimizados comparte entre 65 % y 75 % o aproximadamente 80 % de identidad de secuencia con una secuencia de origen natural o de tipo salvaje (p. ej., una secuencia de ARNm de origen natural o de tipo salvaje que codifica un antígeno de coronavirus).

En algunas modalidades, una secuencia de codones optimizados codifica un antígeno que es tan inmunogénico como, o más inmunogénico que (p. ej., al menos el 10 %, al menos el 20 %, al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 100 %, o al menos el 200 % más), que un antígeno de coronavirus codificado por una secuencia sin codones optimizados.

Cuando se transfectan en células hospedadoras de mamíferos, los ARNm modificados tienen una estabilidad de entre 12-18 horas, o más de 18 horas, p. ej., 24, 36, 48, 60, 72 o más de 72 horas y son capaces de ser expresados por las células hospedadoras de mamíferos.

En algunas modalidades, un ARN optimizado con codones puede ser uno en el que se mejoran los niveles de G/C. El contenido de G/C de las moléculas de ácido nucleico (p. ej., ARNm) puede influir en la estabilidad del ARN. El ARN que tiene una mayor cantidad de residuos de guanina (G) y/o citosina (C) puede ser funcionalmente más estable que el ARN que contiene una gran cantidad de nucleótidos de adenina (A) y timina (T) o uracilo (U). Como ejemplo, el documento WO02/098443 describe una composición farmacéutica que contiene un ARNm estabilizado mediante modificaciones de secuencia en la región traducida. Debido a la degeneración del código genético, las modificaciones funcionan sustituyendo los codones existentes por aquellos que promueven una mayor estabilidad del ARN sin cambiar el aminoácido resultante. El enfoque se limita a codificar regiones del ARN.

Nucleótidos químicamente no modificados

En algunas modalidades, un ARN (p. ej., ARNm) no está modificado químicamente y

comprende los ribonucleótidos estándar que consisten en adenosina, guanosina, citosina y uridina. En algunas modalidades, los nucleótidos y nucleósidos de la presente descripción comprenden residuos de nucleósidos estándar tales como los presentes en el ARN transcrito (p. ej., A, G, C o U). En algunas modalidades, los nucleótidos y nucleósidos de la presente descripción comprenden desoxirribonucleósidos estándar tales como los presentes en el ADN (p. ej., dA, dG, dC o dT).

Modificaciones químicas

Las composiciones de la presente descripción comprenden, en algunas modalidades, un ARN que tiene un marco de lectura abierto que codifica un antígeno de coronavirus, en donde el ácido nucleico comprende nucleótidos y/o nucleósidos que pueden ser estándar (no modificados) o modificados como se conoce en la técnica. En algunas modalidades, los nucleótidos y nucleósidos de la presente descripción comprenden nucleótidos o nucleósidos modificados. Dichos nucleótidos y nucleósidos modificados pueden ser nucleótidos y nucleósidos modificados de origen natural o nucleótidos y nucleósidos modificados de origen no natural. Tales modificaciones pueden incluir aquellas en el azúcar, la cadena principal o la parte de nucleobase del nucleótido y/o nucleósido como se reconoce en la técnica.

En algunas modalidades, un nucleótido o nucleótido modificado de origen natural de la descripción es uno que se conoce o reconoce generalmente en la técnica. Se pueden encontrar ejemplos no limitantes de tales nucleótidos y nucleótidos modificados de origen natural, entre otros, en la base de datos MODOMICS ampliamente reconocida.

En algunas modalidades, un nucleótido o nucleósido modificado de origen no natural de la descripción es uno que se conoce o reconoce generalmente en la técnica. Se pueden encontrar ejemplos no limitantes de tales nucleótidos y nucleósidos modificados de origen no natural, entre otros, en las solicitudes estadounidenses publicadas números PCT/US2012/058519; PCT/US2013/075177; PCT/US2014/058897; PCT/US2014/058891; PCT/US2014/070413; PCT/US2015/36773; PCT/US2015/36759; PCT/US2015/36771; o PCT/IB2017/051367, todas las cuales se incorporan a la presente mediante referencia.

Por lo tanto, los ácidos nucleicos de la descripción (p. ej., ácidos nucleicos de ADN y ácidos nucleicos de ARN, como ácidos nucleicos de ARNm) pueden comprender nucleótidos y nucleósidos estándar, nucleótidos y nucleósidos de origen natural, nucleótidos y

nucleósidos de origen no natural o cualquier combinación de estos.

Los ácidos nucleicos de la descripción (p. ej., ácidos nucleicos de ADN y ácidos nucleicos de ARN, tales como ácidos nucleicos de ARNm), en algunas modalidades, comprenden varios (más de uno) tipos diferentes de nucleótidos y nucleósidos estándar y/o modificados.

- 5 En algunas modalidades, una región particular de un ácido nucleico contiene uno, dos o más (opcionalmente diferentes) tipos de nucleótidos y nucleósidos estándar y/o modificados.

- En algunas modalidades, un ácido nucleico de ARN modificado (p. ej., un ácido nucleico de ARNm modificado), introducido en una célula u organismo, exhibe una degradación
10 reducida en la célula u organismo, respectivamente, en relación con un ácido nucleico no modificado que comprende nucleótidos y nucleósidos estándar.

- En algunas modalidades, un ácido nucleico de ARN modificado (p. ej., un ácido nucleico de ARNm modificado), introducido en una célula u organismo, puede exhibir inmunogenicidad reducida en la célula u organismo, respectivamente (p. ej., una respuesta innata reducida)
15 en relación con un ácido nucleico no modificado que comprende nucleótidos y nucleósidos estándar.

- Los ácidos nucleicos (p. ej., ácidos nucleicos de ARN, tales como ácidos nucleicos de ARNm), en algunas modalidades, comprenden nucleótidos modificados no naturales que se introducen durante la síntesis o postsíntesis de los ácidos nucleicos para lograr las funciones
20 o propiedades deseadas. Las modificaciones pueden estar presentes en enlaces internucleotídicos, bases de purina o pirimidina o azúcares. La modificación se puede introducir con síntesis química o con una enzima polimerasa en el terminal de una cadena o en cualquier otro lugar de la cadena. Cualquiera de las regiones de un ácido nucleico puede modificarse químicamente.

- 25 La presente descripción proporciona nucleósidos y nucleótidos modificados de un ácido nucleico (p. ej., ácidos nucleicos de ARN, tales como ácidos nucleicos de ARNm). Un "nucleósido" se refiere a un compuesto que contiene una molécula de azúcar (p. ej., una pentosa o ribosa) o un derivado de esta en combinación con una base orgánica (p. ej., una purina o pirimidina) o un derivado de esta (también denominada en el presente documento
30 "nucleobase"). Un "nucleótido" se refiere a un nucleósido, incluido un grupo fosfato. Los nucleótidos modificados pueden sintetizarse mediante cualquier método útil, tal como, por ejemplo, químicamente, enzimáticamente o de forma recombinante, para incluir uno o más

nucleósidos modificados o no naturales. Los ácidos nucleicos pueden comprender una región o regiones de nucleósidos enlazados. Estas regiones pueden tener enlaces de estructura variable. Los enlaces pueden ser enlaces fosfodiéster estándar, en cuyo caso los ácidos nucleicos comprenderían regiones de nucleótidos.

- 5 El emparejamiento de bases de nucleótidos modificados abarca no solo los pares de bases estándar adenosina-timina, adenosina-uracilo o guanosina-citosina, sino también pares de bases formados entre nucleótidos y/o nucleótidos modificados que comprenden bases no estándar o modificadas, en donde la disposición de los donantes de enlaces de hidrógeno y los aceptores de enlaces de hidrógeno permiten el enlace de hidrógeno entre una base no estándar y una base estándar o entre dos estructuras de base no estándar complementarias, como, por ejemplo, en aquellos ácidos nucleicos que tienen al menos una modificación química. Un ejemplo de dicho emparejamiento de bases no estándar es el emparejamiento de bases entre el nucleótido modificado inosina y adenina, citosina o uracilo. Puede incorporarse cualquier combinación de base/azúcar o enlazador en los
- 10 ácidos nucleicos de la presente descripción.

En algunas modalidades, las nucleobases modificadas en ácidos nucleicos (p. ej., ácidos nucleicos de ARN, tales como ácidos nucleicos de ARNm) comprenden 1-metil-pseudouridina (m1 ψ), 1-etil-pseudouridina (e1 ψ), 5-metoxi-uridina (mo5U), 5-metil-citidina (m5C) y/o pseudouridina (ψ). En algunas modalidades, las nucleobases modificadas en

20 ácidos nucleicos (p. ej., ácidos nucleicos de ARN, tales como ácidos nucleicos de ARNm) comprenden 5-metoximetiluridina, 5-metiltiouridina, 1-metoximetilpseudouridina, 5-metil citidina y/o 5-metoxitidina. En algunas modalidades, el polirribonucleótido incluye una combinación de al menos dos (p. ej., 2, 3, 4 o más) de cualquiera de las nucleobases modificadas mencionadas anteriormente, incluidas, entre otras, modificaciones químicas.

- 25 En algunas modalidades, un ARNm de la descripción comprende sustituciones de 1-metil-pseudouridina (m1 ψ) en una o más o todas las posiciones de uridina del ácido nucleico.

En algunas modalidades, un ARNm de la descripción comprende sustituciones de 1-metil-pseudouridina (m1 ψ) en una o más o todas las posiciones de uridina del ácido nucleico y sustituciones de 5-metilcitidina en una o más o todas las posiciones de citidina del ácido

30 nucleico.

En algunas modalidades, un ARNm de la descripción comprende sustituciones de pseudouridina (ψ) en una o más o todas las posiciones de uridina del ácido nucleico.

En algunas modalidades, un ARNm de la descripción comprende sustituciones de pseudouridina (ψ) en una o más o todas las posiciones de uridina del ácido nucleico y sustituciones de 5-metil citidina en una o más o todas las posiciones de citidina del ácido nucleico.

- 5 En algunas modalidades, un ARNm de la descripción comprende uridina en una o más o todas las posiciones de uridina del ácido nucleico.

En algunas modalidades, los ARNm se modifican de manera uniforme (p. ej., se modifican completamente, se modifican en toda la secuencia) para una modificación particular. Por ejemplo, un ácido nucleico se puede modificar uniformemente con 1-metil-pseudouridina, lo que significa que todos los residuos de uridina en la secuencia de ARNm se reemplazan con 1-metil-pseudouridina. De manera similar, un ácido nucleico puede modificarse uniformemente para cualquier tipo de residuo de nucleósido presente en la secuencia reemplazándolo con un residuo modificado como los expuestos anteriormente.

- 10 Los ácidos nucleicos de la presente descripción pueden modificarse parcial o totalmente a lo largo de toda la longitud de la molécula. Por ejemplo, uno o más o todos o un tipo dado de nucleótidos (p. ej., purina o pirimidina, o uno cualquiera o más o todos de A, G, U, C) pueden modificarse uniformemente en un ácido nucleico de la descripción, o en una región de secuencia predeterminada de este (p. ej., en el ARNm que incluye o excluye la cola de poli(A)). En algunas modalidades, todos los nucleótidos X en un ácido nucleico de la presente descripción (o en una región de secuencia del mismo) son nucleótidos modificados, en los que X puede ser cualquiera de los nucleótidos A, G, U, C o cualquiera de las combinaciones A+G, A+U, A+C, G+U, G+C, U+C, A+G+U, A+G+C, G+U+C o A+G+C.

- 25 El ácido nucleico puede contener de aproximadamente 1 % a aproximadamente 100 % de nucleótidos modificados (ya sea en relación con el contenido total de nucleótidos o en relación con uno o más tipos de nucleótidos, es decir, uno o más de A, G, U o C) o cualquier porcentaje intermedio (p. ej., del 1 % al 20 %, del 1 % al 25 %, del 1 % al 50 %, del 1 % al 60 %, del 1 % al 70 %, del 1 % al 80 %, del 1 % al 90 %, del 1 % al 95 %, del 10 % al 20 %, del 10 % al 25 %, del 10 % al 50 %, del 10 % al 60 %, del 10 % al 70 %, del 10 % al 80 %, del 10 % al 90 %, del 10 % al 95 %, del 10 % al 100 %, del 20 % al 25 %, del 20 % al 50 %, del 20 % al 60 %, del 20 % al 70 %, del 20 % al 80 %, del 20 % al 90 %, del 20 % al 95 %, del 20 % al 100 %, del 50 % al 60 %, del 50 % al 70 %, del 50 % al 80 %, del 50 % al 90 %, del 50 % al 100 %).

del 50 % al 95 %, del 50 % al 100 %, del 70 % al 80 %, del 70 % al 90 %, del 70 % al 95 %, del 70 % al 100 %, del 80 % al 90 %, del 80 % al 95 %, del 80 % al 100 %, del 90 % al 95 %, del 90 % al 100 % y del 95 % al 100 %). Se entenderá que cualquier porcentaje restante se explica por la presencia de A, G, U o C sin modificar.

- 5 Los ARNm pueden contener un mínimo del 1 % y un máximo del 100 % de nucleótidos modificados, o cualquier porcentaje intermedio, como al menos un 5 % de nucleótidos modificados, al menos un 10 % de nucleótidos modificados, al menos un 25 % de nucleótidos modificados, al menos un 50 % de nucleótidos modificados, al menos 80 % de nucleótidos modificados o al menos 90 % de nucleótidos modificados. Por ejemplo, los
- 10 ácidos nucleicos pueden contener una pirimidina modificada, como un uracilo o una citosina modificados. En algunas modalidades, al menos el 5 %, al menos el 10 %, al menos el 25 %, al menos el 50 %, al menos el 80 %, al menos el 90 % o el 100 % del uracilo en el ácido nucleico se reemplaza con un uracilo modificado (p. ej., un uracilo sustituido en 5). El uracilo modificado se puede reemplazar por un compuesto que tiene una única estructura única, o
- 15 se puede reemplazar por una multiplicidad de compuestos que tienen diferentes estructuras (p. ej., 2, 3, 4 o más estructuras únicas). En algunas modalidades, al menos el 5 %, al menos el 10 %, al menos el 25 %, al menos el 50 %, al menos el 80 %, al menos el 90 % o el 100 % de la citosina en el ácido nucleico se reemplaza con una citosina modificada (p. ej., una citosina sustituida en 5). La citosina modificada se puede reemplazar por un compuesto que
- 20 tiene una sola estructura única, o se puede reemplazar por una multiplicidad de compuestos que tienen diferentes estructuras (p. ej., 2, 3, 4 o más estructuras únicas).

Regiones sin traducir (UTR)

- Los ARNm de la presente descripción pueden comprender una o más regiones o partes que
- 25 actúan o funcionan como una región no traducida. Cuando los ARNm están diseñados para codificar al menos un antígeno de interés, el nucleico puede comprender una o más de estas regiones no traducidas (UTR). Las regiones no traducidas de tipo salvaje de un ácido nucleico se transcriben, pero no se traducen. En el ARNm, la UTR 5' comienza en el sitio de inicio de la transcripción y continúa hasta el codón de inicio, pero no incluye el codón de
- 30 inicio; mientras que, la UTR 3' comienza inmediatamente luego del codón de terminación y continúa hasta la señal de terminación de la transcripción. Existe un creciente conjunto de pruebas acerca de las funciones reguladoras que desempeñan las UTR en términos de

estabilidad y traducción de la molécula de ácido nucleico. Las características reguladoras de una UTR pueden incorporarse en los polinucleótidos de la presente descripción para, entre otras cosas, mejorar la estabilidad de la molécula. Las características específicas también se pueden incorporar para asegurar una regulación a la baja controlada de la transcripción en caso de que estén mal dirigidas a sitios de órganos no deseados. Se conocen y están disponibles en la técnica una variedad de secuencias 5' UTR y 3' UTR.

Una UTR 5' es una región de un ARNm que está directamente corriente arriba (5') del codón de inicio (el primer codón de una transcripción de ARNm traducida por un ribosoma). Una UTR 5' no codifica una proteína (no codifica). Los UTR 5' naturales tienen características que juegan un papel en el inicio de la traducción. Albergan firmas como las secuencias de Kozak, que comúnmente se sabe que están involucradas en el proceso por el cual el ribosoma inicia la traducción de muchos genes. Las secuencias de Kozak tienen el consenso CCR (A/G) CCAUGG (SEQ ID NO: 44), donde R es una purina (adenina o guanina) tres bases cadena arriba del codón de inicio (AUG), que va seguido de otra 'G'. También se sabe que la 5' UTR forma estructuras secundarias que están implicadas en la unión del factor de alargamiento.

En algunas modalidades de la descripción, una 5' UTR es una UTR heteróloga, es decir, es una UTR encontrada en la naturaleza asociada con un ORF diferente. En otra modalidad, una 5' UTR es una UTR sintética, es decir, no se encuentra en la naturaleza. Las UTR sintéticas incluyen las UTR que se han mutado para mejorar sus propiedades, p. ej., las que aumentan la expresión génica y las que son completamente sintéticas. Los ejemplos de 5' UTR incluyen *Xenopus* o α -globina o β -globina derivada de humanos (8278063; 9012219), polipéptido del citocromo b-245 a humano e hidroxesteroide (17b) deshidrogenasa y virus del grabado del tabaco (US8278063, 9012219). También puede utilizarse el gen CMV inmediato-temprano 1 (IE1) (US2014/0206753, WO2013/185069), la secuencia GGGAUCCUACC (SEQ ID NO: 45) (WO2014/144196). En otra modalidad, 5' UTR de un gen TOP es una 5' UTR de un gen TOP que carece del motivo 5' TOP (el tracto de oligopirimidina) (p. ej., WO/2015/101414, WO2015/101415, WO/2015/062738, WO2015/024667, WO2015/024667; elemento 5' UTR derivado del gen de la proteína ribosómica Large 32 (L32) (WO/20151/01414, WO2015/101415, WO/2015/062738), elemento 5' UTR derivado de la 5' UTR de un gen hidroxesteroide (17- β) deshidrogenasa 4 (HSD17B4) (WO2015/024667), o un elemento UTR 5' derivado de la UTR 5' de ATP5A1

(WO2015/024667). En algunas modalidades, se utiliza un sitio de entrada de ribosoma interno (IRES) en lugar de una 5' UTR.

En algunas modalidades, una 5' UTR de la presente descripción comprende una secuencia seleccionada de SEQ ID NO: 2 y SEQ ID NO: 36.

- 5 Una UTR 3' es una región de un ARNm que está directamente corriente abajo (3') del codón de terminación (el codón de una transcripción de ARNm que indica la terminación de la traducción). Una UTR 3' no codifica una proteína (no es codificante). Se sabe que las UTR 3' naturales o de tipo salvaje tienen tramos de adenosinas y uridinas incrustadas en estas. Estas firmas ricas en AU son particularmente frecuentes en genes con altas tasas de
- 10 recambio. En función de sus características de secuencia y propiedades funcionales, los elementos ricos en AU (ARE) se pueden separar en tres clases (Chen et al, 1995): Los ARE de clase I contienen varias copias dispersas de un motivo AUUUA dentro de regiones ricas en U. C-Myc y MyoD contienen ARE de clase I. Los ARE de Clase II poseen dos o más nonámeros UUAUUUA(U/A)(U/A) (SEQ ID NO: 46) superpuestos. Las moléculas que
- 15 contienen este tipo de ARE incluyen GM-CSF y TNF- α . Los ARE de clase III están menos definidos. Estas regiones ricas en U no contienen un motivo AUUUA. c-Jun y Myogenin son dos ejemplos bien estudiados de esta clase. Se sabe que la mayoría de las proteínas que se unen a los ARE desestabilizan al mensajero, mientras que se ha documentado que los miembros de la familia ELAV, sobre todo HuR, aumentan la estabilidad del ARNm. HuR se
- 20 une a ARE de las tres clases. La modificación de los sitios de unión específicos de HuR en UTR 3' de las moléculas de ácido nucleico conducirá a la unión de HuR y, por lo tanto, a la estabilización del mensaje *in vivo*.

- La introducción, eliminación o modificación de elementos ricos en UTR 3' AU (ARE) se puede utilizar para modular la estabilidad de los ácidos nucleicos (p. ej., ARN) de la
- 25 descripción. Cuando se manipulan ácidos nucleicos específicos por ingeniería genética, se pueden introducir una o más copias de un ARE para hacer que los ácidos nucleicos de la descripción sean menos estables y, por lo tanto, reducir la traducción y disminuir la producción de la proteína resultante. Asimismo, los ARE se pueden identificar y eliminar o mutar para aumentar la estabilidad intracelular y así aumentar la traducción y producción de
- 30 la proteína resultante. Los experimentos de transfección se pueden realizar en líneas celulares relevantes, utilizando ácidos nucleicos de la descripción y la producción de proteínas se puede analizar en varios puntos de tiempo luego de la transfección. Por

ejemplo, las células se pueden transfectar con diferentes moléculas de ingeniería ARE y utilizando un kit ELISA para la proteína relevante y la proteína de ensayo producida a las 6 horas, 12 horas, 24 horas, 48 horas y 7 días luego de la transfección.

Las 3' UTR pueden ser heterólogas o sintéticas. Con respecto a las UTR 3', las UTR de globina, incluidas las UTR de β -globina de *Xenopus* y las UTR de β -globina humana, se conocen en la técnica (8278063, 9012219, US2011/0086907). Se ha desarrollado una construcción de β -globina modificada con estabilidad mejorada en algunos tipos de células mediante la clonación de dos β -globina 3' UTR secuenciales humanos cabeza a cola y es bien conocida en la técnica (US2012/0195936, WO2014/071963). Además, también se conocen en la técnica α 2-globina, α 1-globina, UTR y mutantes de estas (WO2015/101415, WO2015/024667). Otras 3' UTR descritas en las construcciones de ARNm en la bibliografía que no es de patentes incluyen CYBA (Ferizi et al., 2015) y albúmina (Thess et al., 2015). Otros ejemplos de 3' UTR incluyen el de la hormona del crecimiento bovina o humana (tipo salvaje o modificado) (WO2013/185069, US2014/0206753, WO2014152774), β -globina de conejo y virus de la hepatitis B (HBV), α -globina 3' UTR y VEEV viral. Las secuencias 3' UTR también se conocen en la técnica. En algunas modalidades, se utiliza la secuencia UUUGAAUU (WO2014/144196). En algunas modalidades, se utilizan 3 UTR de proteína ribosómica humana y de ratón. Otros ejemplos incluyen rps9 3' UTR (WO2015/101414), Figura 4 (WO2015/101415) y albúmina humana 7 (WO2015/101415).

En algunas modalidades, una UTR 3' de la presente descripción comprende una secuencia seleccionada de SEQ ID NO: 4 y SEQ ID NO: 37.

Los expertos en la técnica entenderán que las UTR 5' que son heterólogas o sintéticas pueden utilizarse con cualquier secuencia UTR 3' deseada. Por ejemplo, se puede utilizar una UTR 5' heteróloga con una UTR 3' sintética con una UTR heteróloga de 3'.

Las secuencias que no son UTR también se pueden utilizar como regiones o subregiones dentro de un ácido nucleico. Por ejemplo, pueden incorporarse intrones o partes de secuencias de intrones en regiones de ácido nucleico de la descripción. La incorporación de secuencias intrónicas puede incrementar la producción de proteínas así como los niveles de ácido nucleico.

Las combinaciones de características pueden incluirse en regiones flanqueantes y pueden estar incluidas dentro de otras características. Por ejemplo, el ORF puede estar flanqueado por una UTR 5' que puede contener una fuerte señal de iniciación de la traducción de Kozak

y/o una UTR 3' que puede incluir una secuencia oligo (dT) para la adición de plantilla de una cola poli-A. La 5' UTR puede comprender un primer fragmento de polinucleótido y un segundo fragmento de polinucleótido de este y/o genes diferentes, como las 5' UTR descritas en la publicación de solicitud de patente estadounidense núm. 2010/0293625 y

5 PCT/US2014/069155, que se incorpora a la presente mediante referencia en su totalidad.

Debe entenderse que cualquier UTR de cualquier gen puede incorporarse en las regiones de un ácido nucleico. Además, se pueden utilizar múltiples UTR de tipo salvaje de cualquier gen conocido. También está dentro del alcance de la presente descripción proporcionar UTR artificiales que no sean variantes de regiones de tipo salvaje. Estas UTR o partes de estas

10 pueden colocarse en la misma orientación que en la transcripción de la que fueron seleccionadas o pueden estar alteradas en orientación o ubicación. Por lo tanto, una UTR 5' o 3' puede invertirse, acortarse, alargarse, hacerse con una o más UTR 5' o UTR 3' diferentes. Como se utiliza en este documento, el término "alterado" en lo que se refiere a una secuencia UTR, significa que la UTR se ha cambiado de alguna manera en relación con

15 una secuencia de referencia. Por ejemplo, una UTR 3' o UTR 5' puede alterarse con respecto a una UTR nativa o de tipo salvaje por el cambio en la orientación o ubicación como se enseñó anteriormente o puede alterarse por la inclusión de nucleótidos adicionales, eliminación de nucleótidos, intercambio o transposición de nucleótidos. Cualquiera de estos cambios que produzcan una UTR "alterada" (ya sea 3' o 5') comprende una UTR variante.

En algunas modalidades, se puede utilizar una UTR doble, triple o cuádruple, tal como una UTR 5' o una UTR 3'. Como se utiliza en el presente documento, una UTR "doble" es aquella en la que dos copias de la misma UTR están codificadas en serie o sustancialmente en serie. Por ejemplo, se puede utilizar una doble beta-globina UTR 3' como se describe en la publicación de patente estadounidense 2010/0129877, cuyo contenido se incorpora a la

25 presente mediante referencia en su totalidad.

También está dentro del alcance de la presente descripción tener UTR con patrones. Como se utiliza en el presente documento, "UTR con patrón" son aquellas UTR que reflejan un patrón repetido o alterno, como ABABAB o AABBAABBAABB o ABCABCABC o variantes de estas repetidas una, dos o más de 3 veces. En estos patrones, cada letra, A, B o C

30 representan una UTR diferente a nivel de nucleótidos.

En algunas modalidades, las regiones flanqueantes se seleccionan de una familia de transcritos cuyas proteínas comparten una función, estructura, característica o propiedad

común. Por ejemplo, los polipéptidos de interés pueden pertenecer a una familia de proteínas que se expresan en una célula, tejido particular o en algún momento durante el desarrollo. Las UTR de cualquiera de estos genes pueden intercambiarse por cualquier otra UTR de la misma familia de proteínas o de una diferente para crear un nuevo polinucleótido.

- 5 Como se utiliza en este documento, una "familia de proteínas" se utiliza en el sentido más amplio para referirse a un grupo de dos o más polipéptidos de interés que comparten al menos una función, estructura, característica, localización, origen o patrón de expresión. La región no traducida también puede incluir elementos potenciadores de la traducción (TEE). Como ejemplo no taxativo, el TEE puede incluir los descritos en la Solicitud estadounidense núm. 2009/0226470, que se incorpora a la presente mediante referencia en su totalidad, y los conocidos en la técnica.
- 10

Transcripción de ARN *in vitro*

- El ADNc que codifica los polinucleótidos descritos en el presente documento puede transcribirse utilizando un sistema de transcripción *in vitro* (IVT). La transcripción *in vitro* de ARN se conoce en la técnica y se describe en la publicación internacional WO 2014/152027, que se incorpora a la presente mediante referencia en su totalidad. En algunas modalidades, el ARN de la presente descripción se prepara de conformidad con uno cualquiera o más de los métodos descritos en los documentos WO 2018/053209 y WO 2019/036682, cada uno de los cuales se incorpora a la presente mediante referencia.
- 15
- 20

- En algunas modalidades, la transcripción de ARN se genera utilizando un molde de ADN linealizado no amplificado en una reacción de transcripción *in vitro* para generar la transcripción de ARN. En algunas modalidades, el ADN molde es ADN aislado. En algunas modalidades, el ADN molde es ADNc. En algunas modalidades, el ADNc se forma mediante transcripción inversa de un polinucleótido de RNA, por ejemplo, de modo no taxativo mRNA de coronavirus. En algunas modalidades, las células, p. ej., células bacterianas, p. ej., *E. coli*, p. ej., células DH-1 se transfectan con el molde de ADN plasmídico. En algunas modalidades, las células transfectadas se cultivan para replicar el ADN plasmídico que luego se aísla y purifica. En algunas modalidades, la plantilla de ADN incluye un promotor de ARN polimerasa, p. ej., un promotor T7 ubicado en 5' y unido operativamente al gen de interés.
- 25
- 30
- En algunas modalidades, una plantilla de transcripción *in vitro* codifica una región no traducida (UTR) 5', contiene un marco de lectura abierto y codifica una UTR 3' y una cola

poli(A). La composición de la secuencia de ácido nucleico particular y la longitud de una plantilla de transcripción *in vitro* dependerán del ARNm codificado por la plantilla.

Una "región 5' no traducida" (UTR) se refiere a una región de un ARNm que está directamente corriente arriba (es decir, 5') del codón de inicio (es decir, el primer codón de una transcripción de ARNm traducida por un ribosoma) que no codifica un polipéptido. Cuando se generan transcripciones de ARN, la UTR 5' puede comprender una secuencia promotora. Estas secuencias promotoras son conocidas en la técnica. Debe entenderse que tales secuencias promotoras no estarán presentes en una vacuna de la descripción.

Una "región 3' no traducida" (UTR) se refiere a una región de un ARNm que está directamente corriente abajo (es decir, 3') del codón de parada (es decir, el codón de una transcripción de ARNm que indica una terminación de la traducción) que no codificar un polipéptido.

Un "marco de lectura abierto" es un tramo continuo de ADN que comienza con un codón de inicio (p. ej., metionina (ATG)) y termina con un codón finalizador (p. ej., TAA, TAG o TGA) y codifica un polipéptido.

Una "cola poli(A)" es una región de ARNm que está corriente abajo, p. ej., directamente corriente abajo (es decir, 3'), de la UTR 3' que contiene múltiples adenosina monofosfatos consecutivos. Una cola de poli(A) puede contener de 10 a 300 monofosfatos de adenosina. Por ejemplo, una cola de poli(A) puede contener 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290 o 300 monofosfatos de adenosina. En algunas modalidades, una cola de poli(A) contiene de 50 a 250 monofosfatos de adenosina. En un entorno biológico relevante (p. ej., en células, *in vivo*), la cola poli(A) funciona para proteger el ARNm de la degradación enzimática, p. ej., en el citoplasma, y ayuda en la terminación de la transcripción y/o la exportación y traducción del ARNm desde el núcleo.

En algunas modalidades, un ácido nucleico incluye de 200 a 3000 nucleótidos. Por ejemplo, un ácido nucleico puede incluir 200 a 500, 200 a 1000, 200 a 1500, 200 a 3000, 500 a 1000, 500 a 1500, 500 a 2000, 500 a 3000, 1000 a 1500, 1000 a 2000, 1000 a 3000, 1500 a 3000 o 2000 a 3000 nucleótidos.

Un sistema de transcripción *in vitro* típicamente comprende un amortiguador de transcripción, nucleótidos trifosfatos (NTP), un inhibidor de RNasa y una polimerasa.

Los NTP pueden fabricarse internamente, pueden seleccionarse de un proveedor o pueden

sintetizarse como se describe en el presente documento. Los NTP pueden seleccionarse, de modo no taxativo, a partir de los descritos en el presente documento, incluidos los NTP naturales y no naturales (modificados).

Se puede utilizar cualquier número de ARN polimerasas o variantes en el método de la presente descripción. La polimerasa puede seleccionarse, de modo no taxativo, a partir de una ARN polimerasa de fago, p. ej., una ARN polimerasa T7, una ARN polimerasa T3, una ARN polimerasa SP6 y/o polimerasas mutantes tales como, de modo no taxativo, polimerasas capaces para incorporar ácidos nucleicos modificados y/o nucleótidos modificados, incluidos ácidos nucleicos y/o nucleótidos modificados químicamente. Algunas modalidades excluyen el uso de DNasa.

En algunas modalidades, la transcripción de ARN se recubre mediante un recubrimiento enzimático. En algunas modalidades, el ARN comprende una cubierta terminal 5', por ejemplo, 7mG(5')ppp(5')NImpNp.

15 **Síntesis química**

Síntesis química en fase sólida. Los ácidos nucleicos de la presente descripción pueden fabricarse en su totalidad o en parte utilizando técnicas de fase sólida. La síntesis química de ácidos nucleicos en fase sólida es un método automatizado en donde las moléculas se inmovilizan sobre un soporte sólido y se sintetizan paso a paso en una solución reactiva. La síntesis en fase sólida es útil en la introducción específica de un sitio de modificaciones químicas en las secuencias de ácido nucleico.

Síntesis química en fase líquida. La síntesis de ácidos nucleicos de la presente descripción mediante la adición secuencial de bloques de construcción de monómeros se puede llevar a cabo en una fase líquida.

25 **Combinación de métodos sintéticos.** Cada uno de los métodos sintéticos analizados anteriormente tiene sus propias ventajas y limitaciones. Se han realizado intentos para combinar estos métodos para superar las limitaciones. Tales combinaciones de métodos están dentro del alcance de la presente descripción. El uso de síntesis química en fase sólida o en fase líquida en combinación con la ligadura enzimática proporciona una forma eficiente de generar ácidos nucleicos de cadena larga que no se pueden obtener mediante síntesis química sola.

Ligadura de regiones o subregiones de ácidos nucleicos

También se puede utilizar el ensamblaje de ácidos nucleicos mediante una ligasa. Las ligasas de ADN o ARN promueven la ligadura intermolecular de los extremos 5' y 3' de las cadenas polinucleotídicas mediante la formación de un enlace fosfodiéster. Se pueden preparar ácidos nucleicos tales como polinucleótidos quiméricos y/o ácidos nucleicos circulares mediante ligación de una o más regiones o subregiones. Los fragmentos de ADN pueden unirse mediante una reacción catalizada por ligasa para crear ADN recombinante con diferentes funciones. Dos oligodesoxinucleótidos, uno con un grupo fosforilo 5' y otro con un grupo hidroxilo 3' libre, sirven como sustratos para una ADN ligasa.

Purificación

La purificación de los ácidos nucleicos descritos en este documento puede incluir, de modo no taxativo, limpieza de ácidos nucleicos, garantía de calidad y control de calidad. La limpieza se puede realizar mediante métodos conocidos en la técnica tales como, entre otros, perlas AGENCOURT® (Beckman Coulter Genomics, Danvers, MA), perlas poli-T, sondas de captura de oligo-T LNATM (EXIQON® Inc, Vedbaek, Dinamarca) o métodos de purificación basados en HPLC tales como, entre otros, HPLC de intercambio aniónico fuerte, HPLC de intercambio aniónico débil, HPLC de fase inversa (RP-HPLC) y HPLC de interacción hidrófoba (HIC-HPLC). El término "purificado" cuando se utiliza en relación con un ácido nucleico tal como un "ácido nucleico purificado" se refiere a uno que está separado de al menos un contaminante. Un "contaminante" es cualquier sustancia que hace que otra sea inadecuada, impura o inferior. Por lo tanto, un ácido nucleico purificado (p. ej., ADN y ARN) está presente en una forma o entorno diferente de aquel en el que se encuentra en la naturaleza, o en una forma o entorno diferente al que existía antes de someterlo a un método de tratamiento o purificación.

Se puede realizar una verificación de garantía de calidad y/o control de calidad utilizando métodos tales como, entre otros, electroforesis en gel, absorbancia UV o HPLC analítica. En algunas modalidades, los ácidos nucleicos pueden secuenciarse mediante métodos que incluyen, de modo no taxativo, PCR con transcriptasa inversa.

Cuantificación

En algunas modalidades, los ácidos nucleicos de la presente descripción pueden

cuantificarse en exosomas o cuando se derivan de uno o más fluidos corporales. Los fluidos corporales incluyen sangre periférica, suero, plasma, ascitis, orina, líquido cefalorraquídeo (LCR), esputo, saliva, médula ósea, líquido sinovial, humor acuoso, líquido amniótico, cerumen, leche materna, líquido de lavado bronqueoalveolar, semen, líquido prostático,

5 líquido preseminal o líquido preeyaculatorio, sudor, materia fecal, cabello, lágrimas, líquido quístico, líquido pleural y peritoneal, líquido pericárdico, linfa, quimo, quilo, bilis, líquido intersticial, menstruación, pus, sebo, vómito, secreciones vaginales, secreción mucosa, agua de heces, jugo pancreático, líquidos de lavado de cavidades sinusales, aspirados broncopulmonares, líquido de la cavidad blastocilo y sangre del cordón umbilical.

10 Alternativamente, los exosomas se pueden recuperar de un órgano seleccionado del grupo que consiste en pulmón, corazón, páncreas, estómago, intestino, vejiga, riñón, ovario, testículos, piel, colon, mama, próstata, cerebro, esófago, hígado y placenta.

Los ensayos se pueden realizar utilizando sondas específicas de constructo, citometría, qRT-PCR, PCR en tiempo real, PCR, citometría de flujo, electroforesis, espectrometría de masas o combinaciones de estos, mientras que los exosomas pueden aislarse utilizando

15 métodos inmunohistoquímicos como los métodos de ensayo inmunoabsorbente ligado a enzimas (ELISA). Los exosomas también pueden aislarse mediante cromatografía de exclusión por tamaño, centrifugación en gradiente de densidad, centrifugación diferencial, ultrafiltración de nanomembranas, captura inmunoabsorbente, purificación por afinidad,

20 separación de microfluidos o combinaciones de estos.

Estos métodos brindan al investigador la capacidad de monitorear, en tiempo real, el nivel de ácidos nucleicos restantes o administrados. Esto es posible porque los ácidos nucleicos de la presente descripción, en algunas modalidades, difieren de las formas endógenas debido a las modificaciones estructurales o químicas.

25 En algunas modalidades, el ácido nucleico puede cuantificarse utilizando métodos tales como, de modo no taxativo, espectroscopía ultravioleta visible (UV/Vis). Un ejemplo no taxativo de un espectrómetro UV/Vis es un espectrómetro NANODROP® (ThermoFisher, Waltham, MA). El ácido nucleico cuantificado puede analizarse para determinar si el ácido nucleico puede tener el tamaño adecuado, y comprobar que no se ha producido degradación

30 del ácido nucleico. La degradación del ácido nucleico se puede comprobar mediante métodos como, entre otros, electroforesis en gel de agarosa, métodos de purificación basados en HPLC como, entre otros, HPLC de intercambio aniónico fuerte, HPLC de

intercambio aniónico débil, HPLC de fase inversa (RP HPLC), y HPLC de interacción hidrofóbica (HIC-HPLC), cromatografía líquida-espectrometría de masas (LCMS), electroforesis capilar (CE) y electroforesis capilar en gel (CGE).

5 **Nanopartículas lipídicas (LNP)**

En algunas modalidades, el ARN (p. ej., ARNm) de la descripción se formula en una nanopartícula lipídica (LNP). Las nanopartículas lipídicas comprenden típicamente componentes lipídicos catiónicos ionizables, lípidos no catiónicos, esteroles y PEG junto con la carga de ácido nucleico de interés. Las nanopartículas lipídicas de la descripción se pueden generar utilizando componentes, composiciones y métodos como se conocen generalmente en la técnica, véase por ejemplo PCT/US2016/052352; PCT/US2016/068300; PCT/US2017/037551; PCT/US2015/027400; PCT/US2016/047406; PCT/US2016000129; PCT/US2016/014280; PCT/US2016/014280; PCT/US2017/038426; PCT/US2014/027077; PCT/US2014/055394; PCT/US2016/52117; PCT/US2012/069610; PCT/US2017/027492; PCT/US2016/059575 y PCT/US2016/069491, todos los cuales se incorporan a la presente mediante referencia en su totalidad.

Las vacunas de la presente descripción se formulan típicamente en nanopartículas lipídicas. En algunas modalidades, la nanopartícula lipídica comprende al menos un lípido catiónico ionizable, al menos un lípido no catiónico, al menos un esteroles y/o al menos un lípido modificado con polietilenglicol (PEG).

En algunas modalidades, la nanopartícula lipídica comprende 20-60 % en moles de lípido catiónico ionizable. Por ejemplo, la nanopartícula de lípidos puede comprender 20-50 % en moles, 20-40 % en moles, 20-30 % en moles, 30-60 % en moles, 30-50 % en moles, 30-40 % en moles, 40-60 % en moles, 40-50 % en moles, 50 % en moles o 60 % en moles de lípidos catiónicos ionizables. En algunas modalidades, la nanopartícula de lípidos comprende 20 % en moles, 30 % en moles, 40 % en moles, 50 % en moles o 60 % en moles de lípidos catiónicos ionizables.

En algunas modalidades, la nanopartícula de lípidos comprende 5-25 % en moles de lípidos no catiónicos. Por ejemplo, la nanopartícula lipídica puede comprender 5-20 % en moles, 5-15 % en moles, 5-10 % en moles, 10-25 % en moles, 10-20 % en moles, 10-25 % en moles, 15-25 % en moles, 15-20 % en moles, o 20-25 % en moles de lípidos no catiónicos. En algunas modalidades, la nanopartícula lipídica comprende 5 % en moles, 10 % en moles,

15 % en moles, 20 % en moles o 25 % en moles de lípidos no catiónicos.

En algunas modalidades, la nanopartícula lipídica comprende 25-55 % en moles de esterol.

Por ejemplo, la nanopartícula lipídica puede comprender 25-50 % en moles, 25-45 % en

moles, 25-40 % en moles, 25-35 % en moles, 25-30 % en moles, 30-55 % en moles, 30-

5 50 % en moles, 30-45 % en moles, 30-40 % en moles, 30-35 % en moles, 35-55 % en

moles, 35-50 % en moles, 35-45 % en moles, 35-40 % en moles, 40-55 % en moles, 40-

50 % en moles, 40-45 % en moles, 45-55 % en moles, 45-50 % en moles o 50-55 % en

moles de esterol. En algunas modalidades, la nanopartícula de lípidos comprende 25 % en

moles, 30 % en moles, 35 % en moles, 40 % en moles, 45 % en moles, 50 % en moles o

10 55 % en moles de esterol.

En algunas modalidades, la nanopartícula lipídica comprende 0.5-15 % en moles de lípido

modificado con PEG. Por ejemplo, la nanopartícula lipídica puede comprender 0.5-10 % en

moles, 0.5-5 % en moles, 1-15 % en moles, 1-10 % en moles, 1-5 % en moles, 2-15 % en

moles, 2-10 % en moles, 2-5 % en moles, 5-15 % en moles, 5-10 % en moles o 10-15 % en

15 moles. En algunas modalidades, la nanopartícula lipídica comprende 0.5 % en moles, 1 %

en moles, 2 % en moles, 3 % en moles, 4 % en moles, 5 % en moles, 6 % en moles, 7 % en

moles, 8 % en moles, 9 % en moles, 10 moles %, 11 % en moles, 12 % en moles, 13 % en

moles, 14 % en moles o 15 % en moles de lípidos modificados con PEG.

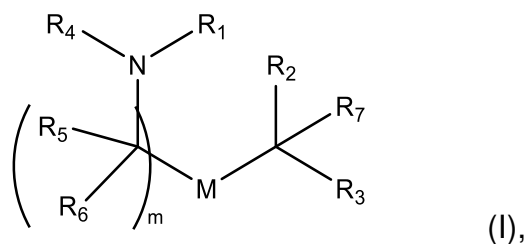
En algunas modalidades, la nanopartícula lipídica comprende 20-60 % en moles de lípido

20 catiónico ionizable, 5-25 % en moles de lípidos no catiónicos, 25-55 % en moles de esterol

y 0.5-15 % en moles de lípidos modificados con PEG.

En algunas modalidades, un lípido catiónico ionizable de la descripción comprende un

compuesto de Fórmula (I):



25 o una sal o isómero de este, en donde:

R₁ se selecciona del grupo que consiste en alquilo C₅₋₃₀, alquenilo C₅₋₂₀, -R*YR'', -YR'' y -R''M'R';

R₂ y R₃ se seleccionan independientemente del grupo que consiste en H, alquilo C₁₋₁₄,

alquenilo C₂₋₁₄, -R*YR", -YR" y -R*OR", o R₂ y R₃, junto con el átomo al que están unidos, forman un heterociclo o carbociclo;

R₄ se selecciona del grupo que consiste en un carbociclo C₃₋₆, -(CH₂)_nQ, -(CH₂)_nCHQR, -CHQR, -CQ(R)₂, y alquilo C₁₋₆ no sustituido, donde Q se selecciona de un carbociclo, heterociclo, -OR, -O (CH₂)_nN(R)₂, -C(O)OR, -OC(O)R, -CX₃, -CX₂H, -CXH₂, -CN, -N(R)₂, -C(O)N(R)₂, -N(R)C(O)R, -N(R)S(O)₂R, -N(R)C(O)N(R)₂, -N(R)C(S)N(R)₂, -N(R)R₈, -O(CH₂)_nOR, -N(R)C(=NR₉)N(R)₂, -N(R)C(=CHR₉)N(R)₂, -OC(O)N(R)₂, -N(R)C(O)O, -N(O)C(O)R, -N(O)S(O)₂R, -N(O)C(O)O, -N(O)C(O)N(R)₂, -N(O)C(S)N(R)₂, -N(O)C(=NR₉)N(R)₂, -N(OR)C(=CHR₉)N(R)₂, -C(=NR₉)N(R)₂, -C(=NR₉)R, -C(O)N(R)OR y -C(R)N(R)₂C(O)OR, y cada n se selecciona independientemente entre 1, 2, 3, 4 y 5;

cada R₅ se selecciona independientemente del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

cada R₆ se selecciona independientemente del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

M y M' se seleccionan independientemente de -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(OR')O-, -S(O)₂-, -SS-, un grupo arilo y un grupo heteroarilo;

R₇ se selecciona del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

R₈ se selecciona del grupo que consiste en C₃₋₆ carbociclo y heterociclo;

R₉ se selecciona del grupo que consiste en H, CN, NO₂, alquilo C₁₋₆, -OR, -S(O)₂R, -S(O)₂N(R)₂, alquenilo C₂₋₆, carbociclo y heterociclo C₃₋₆;

cada R se selecciona independientemente del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

cada R' se selecciona independientemente del grupo que consiste en alquilo C₁₋₁₈, alquenilo C₂₋₁₈, -R*YR", -YR" y H;

cada R" se selecciona independientemente del grupo que consiste en alquilo C₃₋₁₄ y alquenilo C₃₋₁₄;

cada R* se selecciona independientemente del grupo que consiste en alquilo C₁₋₁₂ y alquenilo C₂₋₁₂;

Y es independientemente un carbociclo C₃₋₆;

X se selecciona independientemente del grupo que consiste en F, Cl, Br e I; y

m se selecciona entre 5, 6, 7, 8, 9, 10, 11, 12 y 13.

En algunas modalidades, un subconjunto de compuestos de fórmula (I) incluye aquellos en los que cuando R_4 es $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, $-CHQR$ o $-CQ(R)_2$, entonces (i) Q no es $-N(R)_2$ cuando n es 1, 2, 3, 4 o 5, o (ii) Q no es heterocicloalquilo de 5, 6 o 7 miembros cuando n es 1 o 2.

- 5 En algunas modalidades, otro subconjunto de compuestos de Fórmula (I) incluye aquellos en los que

R_1 se selecciona del grupo que consiste en alquilo C_{5-30} , alquenilo C_{5-20} , $-R^*YR''$, $-YR''$ y $-R''M'R'$;

- 10 R_2 y R_3 se seleccionan independientemente del grupo que consiste en H , alquilo C_{1-14} , alquenilo C_{2-14} , $-R^*YR''$, $-YR''$ y $-R^*OR''$, o R_2 y R_3 , junto con el átomo al que están unidos, forman un heterociclo o carbociclo;

- R_4 se selecciona del grupo que consiste en un carbociclo C_{3-6} , $-(CH_2)_nQ$, $-(CH_2)_nCHQR$, $-CHQR$, $-CQ(R)_2$ y alquilo C_{1-6} no sustituido, donde Q se selecciona entre un carbociclo C_{3-6} , un heteroarilo de 5 a 14 miembros que tiene uno o más heteroátomos seleccionados de N ,
 15 o, y S , $-OR$,
 $-O(CH_2)_nN(R)_2$, $-C(O)O$, $-OC(O)R$, $-CX_3$, $-CX_2H$, $-CXH_2$, $-CN$, $-C(O)N(R)_2$,
 $-N(R)C(O)R$, $-N(R)S(O)_2R$, $-N(R)C(O)N(R)_2$, $-N(R)C(S)N(R)_2$, $-CRN(R)_2C(O)O$, $-N(R)R_8$,
 $-O(CH_2)_nO$, $-N(R)C(=NR_9)N(R)_2$, $-N(R)C(=CHR_9)N(R)_2$, $-OC(O)N(R)_2$, $-N(R)C(O)O$,
 $-N(O)C(O)R$, $-N(O)S(O)_2R$, $-N(O)C(O)O$, $-N(O)C(O)N(R)_2$, $-N(O)C(S)N(R)_2$,
 20 $-N(OR)C(=NR_9)N(R)_2$, $-N(OR)C(=CHR_9)N(R)_2$, $-C(=NR_9)N(R)_2$, $-C(=NR_9)R$, $-C(O)N(R)OR$, y
 un heterocicloalquilo de 5 a 14 miembros que tiene uno o más heteroátomos seleccionados entre N , O y S que está sustituido con uno o más sustituyentes seleccionados de oxo ($=O$), OH , amino, mono- o dialquilamino y alquilo C_{1-3} , y cada n es independientemente seleccionado de 1, 2, 3, 4 y 5;

- 25 cada R_5 se selecciona independientemente del grupo que consiste en alquilo C_{1-3} , alquenilo C_{2-3} y H ;

cada R_6 se selecciona independientemente del grupo que consiste en alquilo C_{1-3} , alquenilo C_{2-3} y H ;

- M y M' se seleccionan independientemente de $-C(O)O-$, $-OC(O)-$, $-C(O)N(R')-$,
 30 $-N(R')C(O)-$, $-C(O)-$, $-C(S)-$, $-C(S)S-$, $-SC(S)-$, $-CH(OH)-$, $-P(O)(O')O-$, $-S(O)_2-$, $-SS-$, un grupo arilo y un grupo heteroarilo;

R_7 se selecciona del grupo que consiste en alquilo C_{1-3} , alquenilo C_{2-3} y H ;

R₈ se selecciona del grupo que consiste en C₃₋₆ carbociclo y heterociclo;

R₉ se selecciona del grupo que consiste en H, CN, NO₂, alquilo C₁₋₆, -OR, -S(O)₂R, -S(O)₂N(R)₂, alquenilo C₂₋₆, carbociclo y heterociclo C₃₋₆;

cada R se selecciona independientemente del grupo que consiste en alquilo C₁₋₃, alquenilo

5 C₂₋₃ y H;

cada R' se selecciona independientemente del grupo que consiste en alquilo C₁₋₁₈, alquenilo C₂₋₁₈, -R*YR'', -YR'' y H;

cada R'' se selecciona independientemente del grupo que consiste en alquilo C₃₋₁₄ y alquenilo C₃₋₁₄;

10 cada R* se selecciona independientemente del grupo que consiste en alquilo C₁₋₁₂ y alquenilo C₂₋₁₂;

cada Y es independientemente un carbociclo C₃₋₆;

cada X se selecciona independientemente del grupo que consiste en F, Cl, Br, e I; y

m se selecciona entre 5, 6, 7, 8, 9, 10, 11, 12 y 13,

15 o sales o isómeros de estos.

En algunas modalidades, otro subconjunto de compuestos de Fórmula (I) incluye aquellos en los que

R₁ se selecciona del grupo que consiste en alquilo C₅₋₃₀, alquenilo C₅₋₂₀, -R*YR'', -YR'' y -R''M'R';

20 R₂ y R₃ se seleccionan independientemente del grupo que consiste en H, alquilo C₁₋₁₄, alquenilo C₂₋₁₄, -R*YR'', -YR'' y -R*OR'', o R₂ y R₃, junto con el átomo al que están unidos, forman un heterociclo o carbociclo;

R₄ se selecciona del grupo que consiste en un carbociclo C₃₋₆, -(CH₂)_nQ, -(CH₂)_nCHQR, -CHQR, -CQ(R)₂ y alquilo C₁₋₆ no sustituido, donde Q se selecciona entre un carbociclo C₃₋₆,

25 un heterociclo de 5 a 14 miembros que tiene uno o más heteroátomos seleccionados de N, o, y S, -OR,

-O(CH₂)_nN(R)₂, -C(O)O, -OC(O)R, -CX₃, -CX₂H, -CXH₂, -CN, -C(O)N(R)₂, -N(R)C(O)R, -N(R)S(O)₂R, -N(R)C(O)N(R)₂, -N(R)C(S)N(R)₂, -CRN(R)₂C(O)OR, -N(R)R₈, -O(CH₂)_nOR, -N(R)C(=NR₉)N(R)₂, -N(R)C(=CHR₉)N(R)₂, -OC(O)N(R)₂, -N(R)C(O)O, -N(O)C(O)R, -

30 N(O)S(O)₂R, -N(O))C(O)O, -N(O)C(O)N(R)₂, -N(O)C(S)N(R)₂, -N(O)C(=NR₉)N(R)₂, -N(OR)C(=CHR₉)N(R)₂, -C(=NR₉)R, -C(O)N(R)OR y -C(=NR₉)N(R)₂, y cada n se selecciona independientemente entre 1, 2, 3, 4 y 5; y cuando Q es un heterociclo de 5 a 14 miembros

y (i) R_4 es $-(CH_2)_nQ$ en el que n es 1 o 2, o (ii) R_4 es $-(CH_2)_nCHQR$ en el que n es 1, o (iii) R_4 es $-CHQR$ y $-CQ(R)_2$, entonces Q es un heteroarilo de 5 a 14 miembros o un heterocicloalquilo de 8 a 14 miembros;

cada R_5 se selecciona independientemente del grupo que consiste en alquilo C_{1-3} , alquenilo

5 C_{2-3} y H;

cada R_6 se selecciona independientemente del grupo que consiste en alquilo C_{1-3} , alquenilo C_{2-3} y H;

M y M' se seleccionan independientemente de $-C(O)O-$, $-OC(O)-$, $-C(O)N(R')-$, $-N(R')C(O)-$, $-C(O)-$, $-C(S)-$, $-C(S)S-$, $-SC(S)-$, $-CH(OH)-$, $-P(O)(O')O-$, $-S(O)_2-$, $-SS-$, un grupo

10 arilo y un grupo heteroarilo;

R_7 se selecciona del grupo que consiste en alquilo C_{1-3} , alquenilo C_{2-3} y H;

R_8 se selecciona del grupo que consiste en carbociclo y heterociclo C_{3-6} ;

R_9 se selecciona del grupo que consiste en H, CN, NO_2 , alquilo C_{1-6} , $-OR$, $-S(O)_2R$, $-S(O)_2N(R)_2$, alquenilo C_{2-6} , carbociclo y heterociclo C_{3-6} ;

15 cada R se selecciona independientemente del grupo que consiste en alquilo C_{1-3} , alquenilo C_{2-3} y H;

cada R' se selecciona independientemente del grupo que consiste en alquilo C_{1-18} , alquenilo C_{2-18} , $-R^*YR''$, $-YR''$ y H;

cada R'' se selecciona independientemente del grupo que consiste en alquilo C_{3-14} y

20 alquenilo C_{3-14} ;

cada R^* se selecciona independientemente del grupo que consiste en alquilo C_{1-12} y alquenilo C_{2-12} ;

cada Y es independientemente un carbociclo C_{3-6} ;

cada X se selecciona independientemente del grupo que consiste en F, Cl, Br, e I; y

25 m se selecciona entre 5, 6, 7, 8, 9, 10, 11, 12 y 13,

o sales o isómeros de estos.

En algunas modalidades, otro subconjunto de compuestos de Fórmula (I) incluye aquellos en los que

R_1 se selecciona del grupo que consiste en alquilo C_{5-30} , alquenilo C_{5-20} , $-R^*YR''$, $-YR''$ y $-R''M'R'$;

30

R_2 y R_3 se seleccionan independientemente del grupo que consiste en H, alquilo C_{1-14} , alquenilo C_{2-14} , $-R^*YR''$, $-YR''$ y $-R^*OR''$, o R_2 y R_3 , junto con el átomo al que están unidos,

forman un heterociclo o carbociclo;

R₄ se selecciona del grupo que consiste en un carbociclo C₃₋₆, -(CH₂)_nQ, -(CH₂)_nCHQR, -CHQR, -CQ(R)₂ y alquilo C₁₋₆ no sustituido, donde Q se selecciona entre un carbociclo C₃₋₆, un heteroarilo de 5 a 14 miembros que tiene uno o más heteroátomos seleccionados de N,

5 o, y S, -OR, -O(CH₂)_nN(R)₂, -C(O)OR, -OC(O)R, -CX₃, -CX₂H, -CXH₂, -CN, -C(O)N(R)₂, -N(R)C(O)R, -N(R)S(O)₂R, -N(R)C(O)N(R)₂, -N(R)C(S)N(R)₂, -CRN(R)₂C(O)OR, -N(R)R₈, -O(CH₂)_nOR, -N(R)C(=NR₉)N(R)₂, -N(R)C(=CHR₉)N(R)₂, -OC(O)N(R)₂, -N(R)C(O)OR, -N(OR)C(O)R, -N(OR)S(O)₂R, -N(OR)C(O)OR, -N(OR)C(O)N(R)₂, -N(OR)C(S)N(R)₂, -N(OR)C(=NR₉)N(R)₂, -N(OR)C(=CHR₉)N(R)₂, -C(=N
10 R₉)R, -C(O)N(R)OR y -C(=NR₉)N(R)₂, y cada n se selecciona independientemente entre 1, 2, 3, 4 y 5;

cada R₅ se selecciona independientemente del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

15 cada R₆ se selecciona independientemente del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

M y M' se seleccionan independientemente de -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(O')O-, -S(O)₂-, -SS-, un grupo arilo y un grupo heteroarilo;

R₇ se selecciona del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

20 R₈ se selecciona del grupo que consiste en carbociclo y heterociclo C₃₋₆;

R₉ se selecciona del grupo que consiste en H, CN, NO₂, alquilo C₁₋₆, -OR, -S(O)₂R, -S(O)₂N(R)₂, alquenilo C₂₋₆, carbociclo y heterociclo C₃₋₆;

cada R se selecciona independientemente del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

25 cada R' se selecciona independientemente del grupo que consiste en alquilo C₁₋₁₈, alquenilo C₂₋₁₈, -R*YR'', -YR'' y H;

cada R'' se selecciona independientemente del grupo que consiste en alquilo C₃₋₁₄ y alquenilo C₃₋₁₄;

30 cada R* se selecciona independientemente del grupo que consiste en alquilo C₁₋₁₂ y alquenilo C₂₋₁₂;

cada Y es independientemente un carbociclo C₃₋₆;

cada X se selecciona independientemente del grupo que consiste en F, Cl, Br, e I; y

m se selecciona entre 5, 6, 7, 8, 9, 10, 11, 12 y 13,

o sales o isómeros de estos.

En algunas modalidades, otro subconjunto de compuestos de Fórmula (I) incluye aquellos en los que

- 5 R_1 se selecciona del grupo que consiste en alquilo C_{5-30} , alquenilo C_{5-20} , $-R^*YR''$, $-YR''$ y $-R''M'R'$;
 R_2 y R_3 se seleccionan independientemente del grupo que consiste en H, alquilo C_{2-14} , alquenilo C_{2-14} , $-R^*YR''$, $-Y-$, y $-R^*OR''$, o R_2 y R_3 , junto con el átomo al que están unidos, forman un heterociclo o carbociclo;
- 10 R_4 es $-(CH_2)_nQ$ o $-(CH_2)_nCHQR$, donde Q es $-N(R)_2$, y n se selecciona entre 3, 4 y 5;
cada R_5 se selecciona independientemente del grupo que consiste en alquilo C_{1-3} , alquenilo C_{2-3} y H;
cada R_6 se selecciona independientemente del grupo que consiste en alquilo C_{1-3} , alquenilo C_{2-3} y H;
- 15 M y M' se seleccionan independientemente de $-C(O)O-$, $-OC(O)-$, $-C(O)N(R')-$, $-N(R')C(O)-$, $-C(O)-$, $-C(S)-$, $-C(S)S-$, $-SC(S)-$, $-CH(OH)-$, $-P(O)(O')O-$, $-S(O)_2-$, $-SS-$, un grupo arilo y un grupo heteroarilo;
 R_7 se selecciona del grupo que consiste en alquilo C_{1-3} , alquenilo C_{2-3} y H;
cada R se selecciona independientemente del grupo que consiste en alquilo C_{1-3} , alquenilo
20 C_{2-3} y H;
cada R' se selecciona independientemente del grupo que consiste en alquilo C_{1-18} , alquenilo C_{2-18} , $-R^*YR''$, $-YR''$ y H;
cada R'' se selecciona independientemente del grupo que consiste en alquilo C_{3-14} y alquenilo C_{3-14} ;
- 25 cada R^* se selecciona independientemente del grupo que consiste de alquilo C_{1-12} y alquenilo C_{1-12} ;
cada Y es independientemente un carbociclo C_{3-6} ;
cada X se selecciona independientemente del grupo que consiste en F, Cl, Br, e I; y
m se selecciona entre 5, 6, 7, 8, 9, 10, 11, 12 y 13,
30 o sales o isómeros de estos.
En algunas modalidades, otro subconjunto de compuestos de Fórmula (I) incluye aquellos en los que

R₁ se selecciona del grupo que consiste en alquilo C₅₋₃₀, alquenilo C₅₋₂₀, -R*YR", -YR" y -R"M'R';

R₂ y R₃ se seleccionan independientemente del grupo que consiste en alquilo C₁₋₁₄, alquenilo C₂₋₁₄, -R*YR", -YR" y -R*OR", o R₂ y R₃, juntos con el átomo al que están unidos, forman un heterociclo o carbociclo;

R₄ se selecciona del grupo que consiste en -(CH₂)_nQ, -(CH₂)_nCHQR, -CHQR y -CQ(R)₂, donde Q es -N(R)₂, y n se selecciona de 1, 2, 3, 4 y 5;

cada R₅ se selecciona independientemente del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

cada R₆ se selecciona independientemente del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

M y M' se seleccionan independientemente de -C(O)O-, -OC(O)-, -C(O)N(R')-, -N(R')C(O)-, -C(O)-, -C(S)-, -C(S)S-, -SC(S)-, -CH(OH)-, -P(O)(O')O-, -S(O)₂-, -SS-, un grupo arilo y un grupo heteroarilo;

R₇ se selecciona del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

cada R se selecciona independientemente del grupo que consiste en alquilo C₁₋₃, alquenilo C₂₋₃ y H;

cada R' se selecciona independientemente del grupo que consiste en alquilo C₁₋₁₈, alquenilo C₂₋₁₈, -R*YR", -YR" y H;

cada R" se selecciona independientemente del grupo que consiste en alquilo C₃₋₁₄ y alquenilo C₃₋₁₄;

cada R* se selecciona independientemente del grupo que consiste de alquilo C₁₋₁₂ y alquenilo C₁₋₁₂;

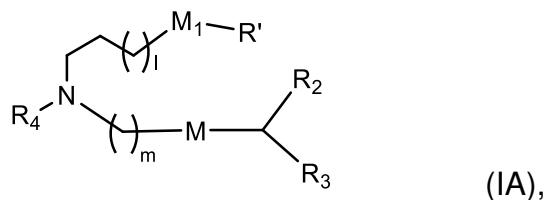
cada Y es independientemente un carbociclo C₃₋₆;

cada X se selecciona independientemente del grupo que consiste en F, Cl, Br, e I; y

m se selecciona entre 5, 6, 7, 8, 9, 10, 11, 12 y 13,

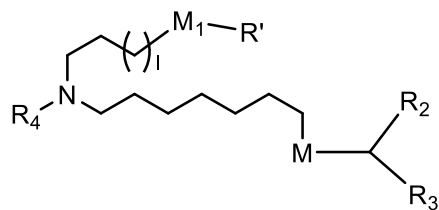
o sales o isómeros de estos.

En algunas modalidades, un subconjunto de compuestos de Fórmula (I) incluye los de Fórmula (IA):



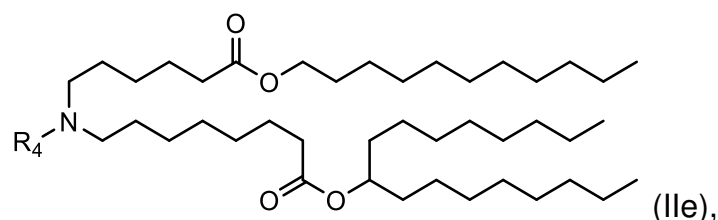
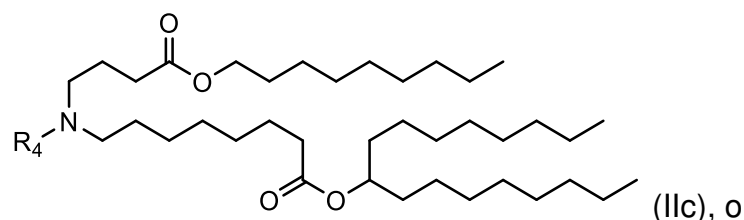
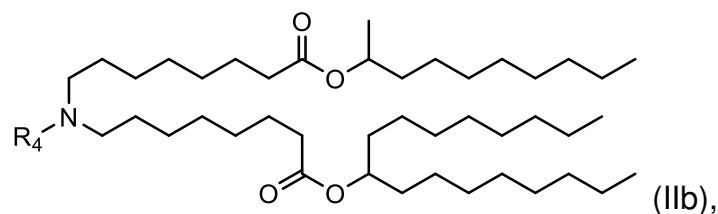
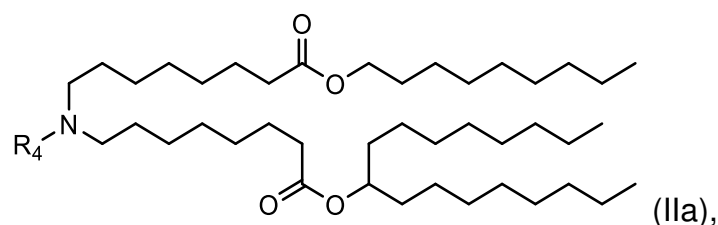
o una sal o isómero de estos, en donde 1 se selecciona entre 1, 2, 3, 4 y 5; m se selecciona entre 5, 6, 7, 8 y 9; M₁ es un enlace o M'; R₄ es alquilo C₁₋₃ no sustituido, o -(CH₂)_nQ, en el que Q es OH, -NHC(S)N(R)₂, -NHC(O)N(R)₂, -N(R)C(O)R, -N(R)S(O)₂R, -N(R)R₈, -NHC(=NR₉)N(R)₂, -NHC(=CHR₉)N(R)₂, -OC(O)N(R)₂, -N(R)C(O)OR, heteroarilo o heterocicloalquilo; M y M' se seleccionan independientemente de -C(O)O-, -OC(O)-, -C(O)N(R')-, -P(O)(OR')O-, -S-S-, un grupo arilo y un grupo heteroarilo; y R₂ y R₃ se seleccionan independientemente del grupo que consiste en H, alquilo C₁₋₁₄ y alqueno C₂₋₁₄.

En algunas modalidades, un subconjunto de compuestos de Fórmula (I) incluye los de Fórmula (II):

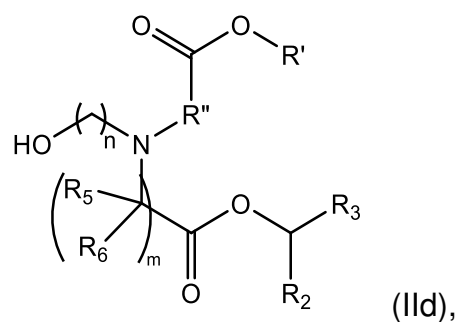


selecciona entre 1, 2, 3, 4 y 5; M₁ es un enlace o M'; R₄ es alquilo C₁₋₃ no sustituido, o -(CH₂)_nQ, en el que n es 2, 3 o 4, y Q es OH, -NHC(S)N(R)₂, -NHC(O)N(R)₂, -N(R)C(O)R, -N(R)S(O)₂R, -N(R)R₈, -NHC(=NR₉)N(R)₂, -NHC(=CHR₉)N(R)₂, -OC(O)N(R)₂, -N(R)C(O)OR, heteroarilo o heterocicloalquilo; M y M' se seleccionan independientemente de -C(O)O-, -OC(O)-, -C(O)N(R')-, -P(O)(OR')O-, -S-S-, un grupo arilo y un grupo heteroarilo; y R₂ y R₃ se seleccionan independientemente del grupo que consiste en H, alquilo C₁₋₁₄ y alqueno C₂₋₁₄.

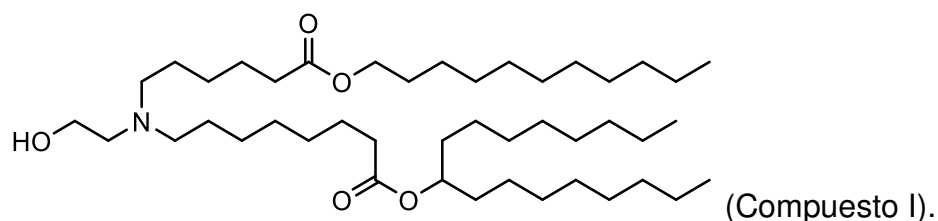
En algunas modalidades, un subconjunto de compuestos de Fórmula (I) incluye los de Fórmula (IIa), (IIb), (IIc) o (Ile):



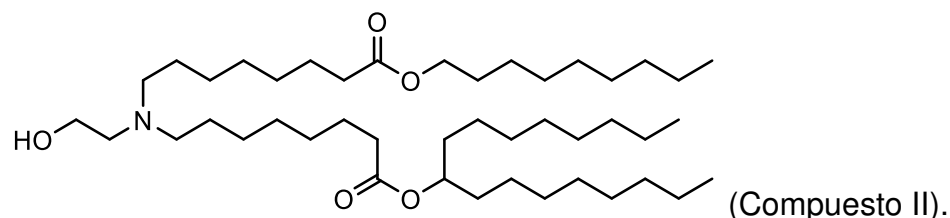
- 5 o una sal o isómero de estos, en el que R₄ es como se describe en este documento.
En algunas modalidades, un subconjunto de compuestos de Fórmula (I) incluye los de Fórmula (IId):



- 10 o una sal o isómero de estos, en donde n es 2, 3 o 4; y m, R', R'' y R₂ a R₆ son como se describen en el presente documento. Por ejemplo, cada uno de R₂ y R₃ se puede seleccionar independientemente del grupo que consiste en alquilo C₅₋₁₄ y alquenilo C₅₋₁₄.
En algunas modalidades, un lípido catiónico ionizable de la descripción comprende un compuesto que tiene la estructura:



En algunas modalidades, un lípido catiónico ionizable de la descripción comprende un compuesto que tiene la estructura:



- 5 En algunas modalidades, un lípido no catiónico de la descripción comprende 1,2-diestearoil-sn-glicero-3-fosfocolina (DSPC), 1,2-dioleoil-sn-glicero-3-fosfoetanolamina (DOPE), 1,2-dilinoil-sn-glicero-3-fosfocolina (DLPC), 1,2-dimiristoil-sn-glicero-fosfocolina (DMPC), 1,2-dioleoil-sn-glicero-3-fosfocolina (DOPC), 1,2-dipalmitoil-sn-glicero-3-fosfocolina (DPPC), 1,2-diundecanoil-sn-glicero-fosfocolina (DUPC), 1-palmitoil-2-oleoil-sn-glicero-3-fosfocolina (POPC), 1,2-di-O-octadecenil-sn-glicero-3-fosfocolina (18:0 Diether PC), 1-oleoil-2-colesterilhemisuccinoil-sn-glicero-3-fosfocolina (OChemSPC), 1-hexadecil-sn-glicero-3-fosfocolina (C16 Lyso PC), 1,2-dilinoil-sn-glicero-3-fosfocolina, 1,2-diarquidonoil-sn-glicero-3-fosfocolina, 1,2-didocosaheptaenoil-sn-glicero-3-fosfocolina, 1,2-difitoil-sn-glicero-3-fosfoetanolamina (ME 16.0 PE), 1,2-diestearoil-sn-glicero-3-fosfoetanolamina, 1,2-dilinoil-sn-glicero-3-fosfoetanolamina, 1,2-dilinoil-sn-glicero-3-fosfoetanolamina, 1,2-diarquidonoil-sn-glicero-3-fosfoetanolamina, 1,2-didocosaheptaenoil-sn-glicero-3-fosfoetanolamina, 1,2-dioleoil-sn-glicero-3-sal sódica de fosfo-rac-(1-glicerol) (DOPG), esfingomielina y mezclas de estos.
- 10
- 15

En algunas modalidades, un lípido modificado con PEG de la descripción comprende una fosfatidiletanolamina modificada con PEG, un ácido fosfatídico modificado con PEG, una ceramida modificada con PEG, una dialquilamina modificada con PEG, un diacilglicerol modificado con PEG, un dialquilmglicerol modificado con PEG y mezclas de estos. En algunas modalidades, el lípido modificado con PEG es DMG-PEG, PEG-c-DOMG (también denominado PEG-DOMG), PEG-DSG y/o PEG-DPG.

20

25 En algunas modalidades, un esteroil de la descripción comprende colesterol, fecosterol,

sitosterol, ergosterol, campesterol, estigmasterol, brasicasterol, tomatidina, ácido ursólico, alfa-tocoferol y mezclas de estos.

En algunas modalidades, un LNP de la descripción comprende un lípido catiónico ionizable del Compuesto 1, en donde el lípido no catiónico es DSPC, el lípido estructural que es colesterol y el lípido PEG es DMG-PEG.

En algunas modalidades, la nanopartícula de lípidos comprende 45 - 55 por ciento en moles (% en moles) de lípidos catiónicos ionizables. Por ejemplo, la nanopartícula de lípidos puede comprender 45, 46, 47, 48, 49, 50, 51, 52, 53, 54 o 55 % en moles de lípidos catiónicos ionizables.

En algunas modalidades, la nanopartícula lipídica comprende 5 - 15 % en moles de DSPC. Por ejemplo, la nanopartícula lipídica puede comprender 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o 15 % en moles de DSPC.

En algunas modalidades, la nanopartícula lipídica comprende 35 - 40 % en moles de colesterol. Por ejemplo, la nanopartícula lipídica puede comprender 35, 36, 37, 38, 39 o 40 % en moles de colesterol.

En algunas modalidades, la nanopartícula lipídica comprende 1 - 2 % en moles de DMG-PEG. Por ejemplo, la nanopartícula lipídica puede comprender 1, 1.5 o 2 % en moles de DMG-PEG.

En algunas modalidades, la nanopartícula lipídica comprende 50 % en moles de lípido catiónico ionizable, 10 % en moles de DSPC, 38.5 % en moles de colesterol y 1.5 % en moles de DMG-PEG.

En algunas modalidades, un LNP de la descripción comprende una relación N:P de aproximadamente 2:1 a aproximadamente 30:1.

En algunas modalidades, un LNP de la descripción comprende una relación N:P de aproximadamente 6:1.

En algunas modalidades, un LNP de la descripción comprende una relación N:P de aproximadamente 3:1.

En algunas modalidades, un LNP de la descripción comprende una relación peso/peso del componente lipídico catiónico ionizable al ARN de aproximadamente 10:1 a aproximadamente 100:1.

En algunas modalidades, un LNP de la descripción comprende una relación peso/peso del componente lipídico catiónico ionizable al ARN de aproximadamente 20:1.

En algunas modalidades, un LNP de la descripción comprende una relación peso/peso del componente lipídico catiónico ionizable al ARN de aproximadamente 10:1.

En algunas modalidades, un LNP de la descripción tiene un diámetro medio de aproximadamente 50 nm a aproximadamente 150 nm.

- 5 En algunas modalidades, un LNP de la descripción tiene un diámetro medio de aproximadamente 70 nm a aproximadamente 120 nm.

Vacunas multivalentes

10 Las composiciones, como se proporciona en el presente documento, pueden incluir ARN o múltiples ARN que codifican dos o más antígenos de la misma especie o de especies diferentes. En algunas modalidades, la composición incluye un ARN o múltiples ARN que codifican dos o más antígenos de coronavirus. En algunas modalidades, el ARN puede codificar 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 o más antígenos de coronavirus.

15 En algunas modalidades, se pueden formular dos o más antígenos que codifican ARN (p. ej., ARNm) diferentes en la misma nanopartícula lipídica. En otras modalidades, se pueden formular dos o más antígenos codificantes de ARN diferentes en nanopartículas lipídicas separadas (cada ARN formulado en una única nanopartícula lipídica). Las nanopartículas lipídicas pueden luego combinarse y administrarse como una única composición de vacuna (p. ej., que comprende múltiples ARN que codifican múltiples antígenos) o pueden
20 administrarse por separado.

Vacunas combinadas

Las composiciones, como se proporciona en el presente documento, pueden incluir un ARN o múltiples ARN que codifican dos o más antígenos de estas o diferentes cepas virales.
25 También se proporcionan aquí vacunas de combinación que incluyen ARN que codifica uno o más coronavirus y uno o más antígenos de un organismo diferente. Por tanto, las vacunas de la presente descripción pueden ser vacunas combinadas que se dirigen a uno o más antígenos de la misma cepa/especie, o uno o más antígenos de diferentes cepas/especies, p. ej., antígenos que inducen inmunidad a organismos que se encuentran en las mismas
30 áreas geográficas donde el riesgo de infección por coronavirus es alto u organismos a los que es probable que una persona esté expuesta cuando se expone a un coronavirus.

Formulaciones farmacéuticas

En el presente documento se proporcionan composiciones (p. ej., composiciones farmacéuticas), métodos, kits y reactivos para la prevención o el tratamiento del coronavirus en humanos y otros mamíferos, por ejemplo. Las composiciones proporcionadas en este documento pueden utilizarse como agentes terapéuticos o profilácticos. Se pueden utilizar en medicina para prevenir y/o tratar una infección por coronavirus.

En algunas modalidades, la vacuna de coronavirus que contiene ARN como se describe en este documento se puede administrar a un sujeto (p. ej., un sujeto mamífero, como un sujeto humano), y los polinucleótidos de ARN se traducen *in vivo* para producir un polipéptido antigénico (antígeno).

Una "cantidad eficaz" de una composición (p. ej., que comprende ARN) se basa, al menos en parte, en el tejido diana, el tipo de célula diana, los medios de administración, las características físicas del ARN (p. ej., la longitud, la composición de nucleótidos y/o extensión de nucleósidos modificados), otros componentes de la vacuna y otros determinantes, tales como edad, peso corporal, altura, sexo y salud general del sujeto. Normalmente, una cantidad eficaz de una composición proporciona una respuesta inmunitaria inducida o potenciada en función de la producción de antígeno en las células del sujeto. En algunas modalidades, una cantidad eficaz de la composición que contiene polinucleótidos de ARN que tienen al menos una modificación química es más eficaz que una composición que contiene un polinucleótido no modificado correspondiente que codifica el mismo antígeno o un antígeno peptídico. El aumento de la producción de antígenos puede demostrarse mediante un aumento de la transfección celular (el porcentaje de células transfectadas con la vacuna de ARN), un aumento de la traducción y/o expresión de proteínas del polinucleótido, una disminución de la degradación del ácido nucleico (como se demuestra, por ejemplo, por una mayor duración de la traducción de la proteína a partir de un polinucleótido modificado) o la respuesta inmune específica de antígeno alterada de la célula huésped.

El término "composición farmacéutica" se refiere a la combinación de un agente activo con un portador, inerte o activo, lo que hace que la composición sea especialmente adecuada para uso diagnóstico o terapéutico *in vivo* o *ex vivo*. Un "portador farmacéuticamente aceptable", luego de administrado a un sujeto, no causa efectos fisiológicos indeseables. El vehículo en la composición farmacéutica debe ser "aceptable" también en el sentido de que

sea compatible con el ingrediente activo y pueda ser capaz de estabilizarlo. Se pueden utilizar uno o más agentes solubilizantes como vehículos farmacéuticos para el suministro de un agente activo. Los ejemplos de un vehículo farmacéuticamente aceptable incluyen, de modo no taxativo, vehículos biocompatibles, adyuvantes, aditivos y diluyentes para lograr una composición utilizable como forma de dosificación. Los ejemplos de otros vehículos incluyen óxido de silicio coloidal, estearato de magnesio, celulosa y lauril sulfato de sodio. Los vehículos y diluyentes farmacéuticos adecuados adicionales, así como las necesidades farmacéuticas para su uso, se describen en Remington's Pharmaceutical Sciences.

En algunas modalidades, las composiciones (que comprenden polinucleótidos y sus polipéptidos codificados) de conformidad con la presente descripción pueden utilizarse para el tratamiento o la prevención de una infección por coronavirus. Una composición puede administrarse de forma profiláctica o terapéutica como parte de un esquema de inmunización activa a individuos sanos o al inicio de la infección durante la fase de incubación o durante la infección activa luego del inicio de los síntomas. En algunas modalidades, la cantidad de ARN proporcionada a una célula, un tejido o un sujeto puede ser una cantidad eficaz para la profilaxis inmunitaria.

Una composición se puede administrar con otros compuestos profilácticos o terapéuticos. Como ejemplo no taxativo, un compuesto profiláctico o terapéutico puede ser un adyuvante o un refuerzo. Como se utiliza en este documento, cuando se hace referencia a una composición profiláctica, como una vacuna, el término "refuerzo" se refiere a una administración adicional de la composición profiláctica (vacuna). Se puede administrar un refuerzo (o vacuna de refuerzo) luego de una administración anterior de la composición profiláctica. El tiempo de administración entre la administración inicial de la composición profiláctica y el refuerzo puede ser, de modo no taxativo, 1 minuto, 2 minutos, 3 minutos, 4 minutos, 5 minutos, 6 minutos, 7 minutos, 8 minutos, 9 minutos, 10 minutos, 15 minutos, 20 minutos, 35 minutos, 40 minutos, 45 minutos, 50 minutos, 55 minutos, 1 hora, 2 horas, 3 horas, 4 horas, 5 horas, 6 horas, 7 horas, 8 horas, 9 horas, 10 horas, 11 horas, 12 horas, 13 horas, 14 horas, 15 horas, 16 horas, 17 horas, 18 horas, 19 horas, 20 horas, 21 horas, 22 horas, 23 horas, 1 día, 36 horas, 2 días, 3 días, 4 días, 5 días, 6 días, 1 semana, 10 días, 2 semanas, 3 semanas, 1 mes, 2 meses, 3 meses, 4 meses, 5 meses, 6 meses, 7 meses, 8 meses, 9 meses, 10 meses, 11 meses, 1 año, 18 meses, 2 años, 3 años, 4 años, 5 años, 6 años, 7 años, 8 años, 9 años, 10 años, 11 años, 12 años, 13 años, 14 años, 15 años, 16

años, 17 años, 18 años, 19 años, 20 años, 25 años, 30 años, 35 años, 40 años, 45 años, 50 años, 55 años, 60 años, 65 años, 70 años, 75 años, 80 años, 85 años, 90 años, 95 años o más de 99 años. En modalidades ilustrativas, el tiempo de administración entre la administración inicial de la composición profiláctica y el refuerzo puede ser, de modo no taxativo, 1 semana, 2 semanas, 3 semanas, 1 mes, 2 meses, 3 meses, 6 meses o 1 año.

En algunas modalidades, una composición puede administrarse por vía intramuscular, intranasal o intradérmica, de manera similar a la administración de vacunas inactivadas conocidas en la técnica.

Se puede utilizar una composición en varios entornos dependiendo de la prevalencia de la infección o del grado o nivel de necesidad médica no satisfecha. Como ejemplo no taxativo, las vacunas de ARN se pueden utilizar para tratar y/o prevenir una variedad de enfermedades infecciosas. Las vacunas de ARN tienen propiedades superiores en cuanto a que producen títulos de anticuerpos mucho mayores, mejor neutralizan la inmunidad, producen respuestas inmunitarias más duraderas y/o producen respuestas antes que las vacunas disponibles comercialmente.

En el presente documento se proporcionan composiciones farmacéuticas que incluyen ARN y/o complejos opcionalmente en combinación con uno o más excipientes farmacéuticamente aceptables.

El ARN se puede formular o administrar solo o junto con uno o más de otros componentes.

Por ejemplo, una composición inmunizante puede comprender otros componentes que incluyen, de modo no taxativo, adyuvantes.

En algunas modalidades, una composición inmunizante no incluye un adyuvante (están libres de adyuvantes).

Un ARN se puede formular o administrar en combinación con uno o más excipientes farmacéuticamente aceptables. En algunas modalidades, las composiciones de vacuna comprenden al menos una sustancia activa adicional, tal como, por ejemplo, una sustancia terapéuticamente activa, una sustancia profilácticamente activa o una combinación de ambas. Las composiciones de vacuna pueden ser estériles, libres de pirógenos o tanto estériles como libres de pirógenos. Se pueden encontrar consideraciones generales en la formulación y/o fabricación de agentes farmacéuticos, tales como composiciones de vacunas, por ejemplo, en Remington: The Science and Practice of Pharmacy 21st ed., Lippincott Williams & Wilkins, 2005 (incorporado a la presente mediante referencia en su

totalidad).

En algunas modalidades, se administra una composición inmunizante a seres humanos, pacientes o sujetos humanos. Para los propósitos de la presente descripción, la frase "ingrediente activo" generalmente se refiere a las vacunas de ARN o los polinucleótidos contenidos en ellas, por ejemplo, polinucleótidos de ARN (p. ej., polinucleótidos de ARNm) que codifican antígenos.

Las formulaciones de las composiciones de vacuna descritas en el presente documento pueden prepararse mediante cualquier método conocido o desarrollado de aquí en adelante en la técnica de la farmacología. En general, tales métodos preparatorios incluyen el paso de asociar el ingrediente activo (p. ej., polinucleótido de ARNm) con un excipiente y/o uno o más de otros ingredientes accesorios, y luego, si es necesario y/o deseable, dividir, dar forma y/o empaquetar el producto en una unidad de dosis única o múltiple deseada.

Las cantidades relativas del ingrediente activo, el excipiente farmacéuticamente aceptable y/o cualquier ingrediente adicional en una composición farmacéutica de conformidad con la descripción variarán, dependiendo de la identidad, tamaño y/o condición del sujeto tratado y además dependiendo de la vía por la que se administrará la composición. A modo ilustrativo, la composición puede comprender entre 0.1 % y 100 %, p. ej., entre 0.5 y 50 %, entre 1-30 %, entre 5-80 %, al menos 80 % (p/p) de ingrediente activo.

En algunas modalidades, se formula un ARN utilizando uno o más excipientes para: (1) aumentar la estabilidad; (2) aumentar la transfección celular; (3) permitir la liberación sostenida o retardada (p. ej., de una formulación de depósito); (4) alterar la biodistribución (p. ej., dirigirse a tejidos o tipos celulares específicos); (5) aumentar la traducción de la proteína codificada *in vivo*; y/o (6) alterar el perfil de liberación de la proteína codificada (antígeno) *in vivo*. Además de los excipientes tradicionales tales como todos y cada uno de los disolventes, medios de dispersión, diluyentes u otros vehículos líquidos, adyuvantes de dispersión o suspensión, agentes tensioactivos, agentes isotónicos, agentes espesantes o emulsionantes, conservantes, excipientes pueden incluir, sin limitación, lípidoides, liposomas, nanopartículas lipídicas, polímeros, lipoplejos, nanopartículas núcleo-cap, péptidos, proteínas, células transfectadas con el ARN (p. ej., para trasplante en un sujeto), hialuronidasa, imitadores de nanopartículas y combinaciones de los mismos.

Dosificación/Administración

En el presente documento se proporcionan composiciones inmunizantes (p. ej., vacunas de ARN), métodos, kits y reactivos para la prevención y/o el tratamiento de la infección por coronavirus en humanos y otros mamíferos. Las composiciones inmunizantes se pueden utilizar como agentes terapéuticos o profilácticos. En algunas modalidades, las composiciones inmunizantes se utilizan para proporcionar protección profiláctica frente a la infección por coronavirus. En algunas modalidades, las composiciones inmunizantes se utilizan para tratar una infección por coronavirus. En algunas modalidades, se utilizan composiciones inmunizantes en el cebado de células efectoras inmunes, por ejemplo, para activar células mononucleares de sangre periférica (PBMC) *ex vivo*, que luego se infunden (se vuelven a infundir) en un sujeto.

Un sujeto puede ser cualquier mamífero, incluidos primates no humanos y sujetos humanos. Normalmente, un sujeto es un sujeto humano.

En algunas modalidades, se administra una composición inmunizante (p. ej., ARN una vacuna) a un sujeto (p. ej., un sujeto mamífero, tal como un sujeto humano) en una cantidad eficaz para inducir una respuesta inmune específica de antígeno. El ARN que codifica el antígeno del coronavirus se expresa y traduce *in vivo* para producir el antígeno, que luego estimula una respuesta inmune en el sujeto.

La protección profiláctica contra un coronavirus se puede lograr luego de la administración de una composición inmunizante (p. ej., una vacuna de ARN) de la presente descripción.

Las composiciones inmunizantes pueden administrarse una, dos, tres veces, cuatro veces o más, pero es probable que sea suficiente administrar la vacuna una vez (opcionalmente seguida de un único refuerzo). Es posible, aunque menos deseable, administrar una composición inmunizante a un individuo infectado para lograr una respuesta terapéutica. Es posible que sea necesario ajustar la dosis en consecuencia.

En aspectos de la presente descripción se proporciona un método para provocar una respuesta inmune en un sujeto contra un antígeno de coronavirus (o múltiples antígenos). En algunas modalidades, un método implica administrar al sujeto una composición inmunizante que comprende un ARN (p. ej., ARNm) que tiene un marco de lectura abierto que codifica un antígeno de coronavirus, induciendo así en el sujeto una respuesta inmune específica al antígeno de coronavirus, en donde el antígeno el título de anticuerpos en el sujeto aumenta luego de la vacunación en relación con el título de anticuerpos antiantígeno en un sujeto vacunado con una dosis profilácticamente eficaz de una vacuna tradicional

contra el antígeno. Un "anticuerpo antiantígeno" es un anticuerpo en suero que se une específicamente al antígeno.

Una dosis profilácticamente eficaz es una dosis eficaz que previene la infección por el virus a un nivel clínicamente aceptable. En algunas modalidades, la dosis eficaz es una dosis enumerada en un prospecto de la vacuna. Una vacuna tradicional, como se utiliza en este documento, se refiere a una vacuna distinta de las vacunas de ARNm de la presente descripción. Por ejemplo, una vacuna tradicional incluye, entre otras, vacunas de microorganismos vivos, vacunas de microorganismos muertos, vacunas de subunidades, vacunas de antígenos proteicos, vacunas de ADN, vacunas de partículas similares a virus (VLP), etc. En modalidades ilustrativas, una vacuna tradicional es una vacuna que ha obtenido la aprobación regulatoria y/o está registrada por un organismo regulador de medicamentos nacional, por ejemplo, la Administración de Alimentos y Medicamentos (FDA) en los Estados Unidos o la Agencia Europea de Medicamentos (EMA).

En algunas modalidades, el título de anticuerpo antiantígeno en el sujeto aumenta de 1 log a 10 log luego de la vacunación en relación con el título de anticuerpo antiantígeno en un sujeto vacunado con una dosis profilácticamente eficaz de una vacuna tradicional contra el coronavirus o un sujeto no vacunado. En algunas modalidades, el título de anticuerpo antiantígeno en el sujeto aumenta 1 log, 2 log, 3 log, 4 log, 5 log o 10 log luego de la vacunación en relación con el título de anticuerpo antiantígeno en un sujeto vacunado con una dosis de fármaco profilácticamente eficaz de una vacuna tradicional contra el coronavirus o un sujeto no vacunado.

En otros aspectos de la descripción se proporciona un método para provocar una respuesta inmune en un sujeto contra un coronavirus. El método implica administrar al sujeto una composición inmunizante (p. ej., una vacuna de ARN) que comprende un polinucleótido de ARN que comprende un marco de lectura abierto que codifica un antígeno de coronavirus, induciendo así en el sujeto una respuesta inmune específica al coronavirus, en donde la respuesta inmune en el sujeto es equivalente a una respuesta inmune en un sujeto vacunado con una vacuna tradicional contra el coronavirus de 2 a 100 veces el nivel de dosificación en relación con la composición inmunizante.

En algunas modalidades, la respuesta inmune en el sujeto es equivalente a una respuesta inmune en un sujeto vacunado con una vacuna tradicional al doble del nivel de dosificación en relación con una composición inmunizante de la presente descripción. En algunas

modalidades, la respuesta inmune en el sujeto es equivalente a una respuesta inmune en un sujeto vacunado con una vacuna tradicional a tres veces el nivel de dosificación en relación con una composición inmunizante de la presente descripción. En algunas modalidades, la respuesta inmune en el sujeto es equivalente a una respuesta inmune en un sujeto vacunado con una vacuna tradicional a 4 veces, 5 veces, 10 veces, 50 veces o 100 veces el nivel de dosificación en relación con una composición inmunizante del presente descripción. En algunas modalidades, la respuesta inmune en el sujeto es equivalente a una respuesta inmune en un sujeto vacunado con una vacuna tradicional de 10 a 1000 veces el nivel de dosificación en relación con una composición inmunizante de la presente descripción. En algunas modalidades, la respuesta inmune en el sujeto es equivalente a una respuesta inmune en un sujeto vacunado con una vacuna tradicional de 100 a 1000 veces el nivel de dosificación en relación con una composición inmunizante de la presente descripción.

En otras modalidades, la respuesta inmune se evalúa determinando el título de anticuerpo [proteína] en el sujeto. En otras modalidades, se prueba la capacidad del suero o del anticuerpo de un sujeto inmunizado para determinar su capacidad para neutralizar la captación viral o reducir la transformación por coronavirus de los linfocitos B humanos. En otras modalidades, la capacidad para promover una respuesta o respuestas robustas de células T se mide utilizando técnicas reconocidas en la técnica.

Otros aspectos de la descripción proporcionan métodos para provocar una respuesta inmune en un sujeto contra un coronavirus administrando al sujeto una composición inmunizante (p. ej., una vacuna de ARN) que comprende un ARN que tiene un marco de lectura abierto que codifica un antígeno de coronavirus, induciendo así en el sujeto una respuesta inmune específica al antígeno del coronavirus, en donde la respuesta inmune en el sujeto se induce de 2 días a 10 semanas antes en relación con una respuesta inmune inducida en un sujeto vacunado con una dosis profilácticamente eficaz de una vacuna tradicional contra el coronavirus. En algunas modalidades, la respuesta inmune en el sujeto se induce en un sujeto vacunado con una dosis profilácticamente eficaz de una vacuna tradicional de 2 a 100 veces el nivel de dosificación en relación con una composición inmunizante de la presente descripción.

En algunas modalidades, la respuesta inmune en el sujeto se induce 2 días, 3 días, 1 semana, 2 semanas, 3 semanas, 5 semanas o 10 semanas antes en relación con una

respuesta inmune inducida en un sujeto vacunado con una dosis de fármaco profilácticamente eficaz de una vacuna tradicional.

También se proporcionan en este documento métodos para provocar una respuesta inmune en un sujeto contra un coronavirus administrando al sujeto un ARN que tiene un marco de lectura abierto que codifica un primer antígeno, en donde el ARN no incluye un elemento de estabilización y en donde un adyuvante no se formula ni administra conjuntamente con la vacuna.

Una composición inmunizante (p. ej., una vacuna de ARN) puede administrarse por cualquier vía que dé como resultado un resultado terapéuticamente eficaz. Estos incluyen, de modo no taxativo, la administración intradérmica, intramuscular, intranasal y/o subcutánea. La presente descripción proporciona métodos que comprenden administrar vacunas de ARN a un sujeto que las necesite. La cantidad exacta requerida variará de un sujeto a otro, en función de la especie, edad y estado general del sujeto, la gravedad de la enfermedad, la composición particular, su modo de administración, su modo de actividad y similares. El ARN se formula típicamente en forma de unidad de dosificación para facilitar la administración y uniformidad de la dosificación. Sin embargo, se entenderá que el médico tratante puede decidir el uso diario total del ARN dentro del alcance del buen juicio médico. El nivel de dosis de formación de imágenes específico terapéuticamente eficaz, profilácticamente eficaz o apropiado para cualquier paciente particular dependerá de una variedad de factores que incluyen el trastorno a tratar y la gravedad del trastorno; la actividad del compuesto específico empleado; la composición específica empleada; la edad, peso corporal, salud general, sexo y dieta del paciente; el momento de administración, la vía de administración y la velocidad de excreción del compuesto específico empleado; la duración del tratamiento; medicamentos utilizados en combinación o coincidentemente con el compuesto específico empleado; y factores similares bien conocidos en las técnicas médicas.

La cantidad eficaz de ARN, como se proporciona en este documento, puede ser tan baja como 20 µg, administrada, por ejemplo, como una dosis única o como dos dosis de 10 µg. En algunas modalidades, la cantidad eficaz es una dosis total de 20 µg-300 µg o 25 µg-300 µg. Por ejemplo, la cantidad eficaz puede ser una dosis total de 20 µg, 25 µg, 30 µg, 35 µg, 40 µg, 45 µg, 50 µg, 55 µg, 60 µg, 65 µg, 70 µg, 75 µg, 80 µg, 85 µg, 90 µg, 95 µg, 100 µg, 110 µg, 120 µg, 130 µg, 140 µg, 150 µg, 160 µg, 170 µg, 180 µg, 190 µg, 200 µg, 250 µg o

300 µg. En algunas modalidades, la cantidad eficaz es una dosis total de 20 µg. En algunas modalidades, la cantidad eficaz es una dosis total de 25 µg. En algunas modalidades, la cantidad eficaz es una dosis total de 75 µg. En algunas modalidades, la cantidad eficaz es una dosis total de 150 µg. En algunas modalidades, la cantidad eficaz es una dosis total de 300 µg.

El ARN descrito en este documento se puede formular en una forma de dosificación descrita en este documento, tal como intranasal, intratraqueal o inyectable (p. ej., intravenosa, intraocular, intravítrea, intramuscular, intradérmica, intracardiaca, intraperitoneal y subcutánea).

Eficacia de la vacuna

Algunos aspectos de la presente descripción proporcionan formulaciones de las composiciones inmunizantes (p. ej., vacunas de ARN), en donde el ARN se formula en una cantidad eficaz para producir una respuesta inmune específica de antígeno en un sujeto (p. ej., producción de anticuerpos específicos para un antígeno de coronavirus). "Una cantidad eficaz" es una dosis de ARN eficaz para producir una respuesta inmunitaria específica de antígeno. También se proporcionan en este documento métodos para inducir una respuesta inmune específica de antígeno en un sujeto.

Como se utiliza en este documento, una respuesta inmune a una vacuna o LNP de la presente descripción es el desarrollo en un sujeto de una respuesta inmune humoral y/o celular a una (una o más) proteínas de coronavirus presentes en la vacuna. Para los propósitos de la presente descripción, una respuesta inmune "humoral" se refiere a una respuesta inmune mediada por moléculas de anticuerpos, que incluyen, p. ej., moléculas secretoras (IgA) o IgG, mientras que una respuesta inmune "celular" es una mediada por linfocitos T (p. ej., linfocitos T CD4 + auxiliares y/o CD8 + (p. ej., CTL) y/u otros glóbulos blancos. Un aspecto importante de la inmunidad celular implica una respuesta específica de antígeno por las células T citolíticas (CTL). Los CTL tienen especificidad por antígenos peptídicos que se presentan en asociación con proteínas codificadas por el complejo principal de histocompatibilidad (MHC) y expresadas en la superficie de las células. Los CTL ayudan a inducir y promover la destrucción de microbios intracelulares o la lisis de células infectadas con dichos microbios. Otro aspecto de la inmunidad celular implica una respuesta específica de antígeno por parte de las células T colaboradoras. Las células T colaboradoras

actúan para ayudar a estimular la función y enfocan la actividad de las células efectoras inespecíficas contra las células que presentan antígenos peptídicos en asociación con moléculas MHC en su superficie. Una respuesta inmune celular también conduce a la producción de citocinas, quimiocinas y otras moléculas similares producidas por células T activadas y/u otros glóbulos blancos, incluidos los derivados de células T CD4+ y CD8+.

En algunas modalidades, la respuesta inmune específica de antígeno se caracteriza midiendo un título de anticuerpo de antígeno anticoronavirus producido en un sujeto al que se le administra una composición inmunizante como se proporciona en la presente descripción. Un título de anticuerpo es una medida de la cantidad de anticuerpos dentro de un sujeto, por ejemplo, anticuerpos que son específicos para un antígeno particular o epítipo de un antígeno. El título de anticuerpos se expresa típicamente como la inversa de la mayor dilución que proporciona un resultado positivo. El ensayo inmunoabsorbente ligado a enzimas (ELISA) es un ensayo común para determinar los títulos de anticuerpos, por ejemplo.

En algunas modalidades, se utiliza un título de anticuerpos para evaluar si un sujeto ha tenido una infección o para determinar si se requieren inmunizaciones. En algunas modalidades, se utiliza un título de anticuerpos para determinar la fuerza de una respuesta autoinmune, para determinar si se necesita una inmunización de refuerzo, para determinar si una vacuna previa fue efectiva e identificar cualquier infección reciente o previa. De conformidad con la presente descripción, puede utilizarse un título de anticuerpos para determinar la fuerza de una respuesta inmune inducida en un sujeto por una composición inmunizante (p. ej., vacuna de ARN).

En algunas modalidades, el título de anticuerpo antigénico anticoronavirus producido en un sujeto aumenta en al menos 1 log con respecto a un control. Por ejemplo, el título de anticuerpo de antígeno anticoronavirus producido en un sujeto puede incrementarse en al menos 1.5, al menos 2, al menos 2.5 o al menos 3 log con respecto a un control. En algunas modalidades, el título de anticuerpos antigénicos anticoronavirus producido en el sujeto aumenta en 1, 1.5, 2, 2.5 o 3 log con respecto a un control. En algunas modalidades, el título de anticuerpo de antígeno anticoronavirus producido en el sujeto aumenta en 1-3 log con respecto a un control. Por ejemplo, el título de anticuerpo de antígeno anticoronavirus producido en un sujeto puede incrementarse en 1-1.5; 1-2; 1-2.5; 1-3; 1.5-2; 1.5-2.5; 1.5-3; 2-2.5; 2-3; o 2.5-3 log relativo a un control.

En algunas modalidades, el título de anticuerpo de antígeno anticoronavirus producido en un sujeto aumenta al menos 2 veces con respecto a un control. Por ejemplo, el título de anticuerpos del antígeno n anticoronavirus producido en un sujeto puede incrementarse al menos 3 veces, al menos 4 veces, al menos 5 veces, al menos 6 veces, al menos 7 veces, al menos 8 veces, al menos 9 veces, o al menos 10 veces en relación con un control. En algunas modalidades, el título de anticuerpo de antígeno anticoronavirus producido en el sujeto aumenta 2, 3, 4, 5, 6, 7, 8, 9 o 10 veces con respecto a un control. En algunas modalidades, el título de anticuerpo de antígeno anticoronavirus producido en un sujeto aumenta de 2-10 veces con respecto a un control. Por ejemplo, el título de anticuerpos de antígeno anticoronavirus producido en un sujeto puede incrementarse 2-10, 2-9, 2-8, 2-7, 2-6, 2-5, 2-4, 2-3, 3-10, 3-9, 3-8, 3-7, 3-6, 3-5, 3-4, 4-10, 4-9, 4-8, 4-7, 4-6, 4-5, 5-10, 5-9, 5-8, 5-7, 5-6, 6-10, 6-9, 6-8, 6-7, 7-10, 7-9, 7-8, 8-10, 8-9 o 9-10 veces en relación con un control. En algunas modalidades, una respuesta inmune específica de antígeno se mide como una relación de título de media geométrica (GMT), denominada relación de media geométrica (GMR), de títulos de anticuerpos neutralizantes en suero frente a coronavirus. Un título medio geométrico (GMT) es el título promedio de anticuerpos para un grupo de sujetos calculado multiplicando todos los valores y tomando la raíz *enésima* del número, donde *n* es el número de sujetos con datos disponibles.

Un control, en algunas modalidades, es un título de anticuerpo de antígeno anticoronavirus producido en un sujeto al que no se le ha administrado una composición inmunizante (p. ej., vacuna de ARN). En algunas modalidades, un control es un título de anticuerpo de antígeno anticoronavirus producido en un sujeto al que se le administra una vacuna de proteína recombinante o purificada. Las vacunas de proteínas recombinantes incluyen típicamente antígenos de proteínas que se han producido en un sistema de expresión heterólogo (p. ej., bacterias o levaduras) o se han purificado a partir de grandes cantidades del organismo patógeno.

En algunas modalidades, la capacidad de una composición inmunizante (p. ej., vacuna de ARN) para ser eficaz se mide en un modelo murino. Por ejemplo, puede administrarse una composición inmunizante a un modelo murino y ensayarse el modelo murino para la inducción de títulos de anticuerpos neutralizantes. También pueden utilizarse estudios de provocación viral para evaluar la eficacia de una vacuna de la presente descripción. Por ejemplo, se puede administrar una composición inmunizante a un modelo murino, el modelo

murino expuesto al virus y el modelo murino ensayado para supervivencia y/o respuesta inmune (p. ej., respuesta de anticuerpos neutralizantes, respuesta de células T (p. ej., respuesta de citocinas)).

En algunas modalidades, una cantidad eficaz de una composición inmunizante (p. ej., vacuna de ARN) es una dosis que se reduce en comparación con la dosis estándar de cuidado de una vacuna de proteína recombinante. Un "estándar de atención", como se proporciona en este documento, se refiere a una guía de tratamiento médico o psicológico y puede ser general o específico. El "estándar de atención" especifica el tratamiento apropiado basado en evidencia científica y la colaboración entre los profesionales médicos involucrados en el tratamiento de una condición determinada. Es el proceso de diagnóstico y tratamiento que debe seguir un médico/especialista para un determinado tipo de paciente, enfermedad o circunstancia clínica. Una "dosis estándar de atención", como se proporciona en este documento, se refiere a la dosis de una vacuna de proteína recombinante o purificada, o una vacuna viva atenuada o inactivada, o una vacuna VLP, que un médico/especialista u otro profesional médico administraría a un sujeto a tratar o prevenir la infección por coronavirus o una afección relacionada, mientras sigue las pautas de atención estándar para tratar o prevenir la infección por coronavirus o una afección relacionada.

En algunas modalidades, el título de anticuerpo de antígeno anticoronavirus producido en un sujeto al que se le administra una cantidad eficaz de una composición inmunizante es equivalente a un título de anticuerpo de antígeno anticoronavirus producido en un sujeto de control al que se le administra una dosis estándar de cuidado de una vacuna proteica recombinante o purificada, o una vacuna viva atenuada o inactivada, o una vacuna VLP.

La eficacia de la vacuna puede evaluarse utilizando análisis estándar (véase, p. ej., Weinberg et al., J Infect Dis. 1 de junio de 2010; 201(11):1607-10). Por ejemplo, la eficacia de la vacuna se puede medir mediante ensayos clínicos controlados aleatorios, con enmascaramiento doble. La eficacia de la vacuna puede expresarse como una reducción proporcional en la tasa de ataque de la enfermedad (RA) entre las cohortes del estudio no vacunadas (ARU) y vacunadas (ARV) y puede calcularse a partir del riesgo relativo (RR) de enfermedad entre el grupo vacunado mediante las siguientes fórmulas:

Eficacia = $(ARU - ARV)/ARU \times 100$; y

Eficacia = $(1 - RR) \times 100$.

Asimismo, la eficacia de la vacuna puede evaluarse mediante análisis estándar (véase, p.

ej., Weinberg et al., J Infect Dis. 1 de junio de 2010; 201(11):1607-10). La eficacia de la vacuna es una evaluación de cómo una vacuna (que ya puede haber demostrado tener una alta eficacia) reduce la enfermedad en una población. Esta medida puede evaluar el balance neto de beneficios y efectos adversos de un programa de vacunación, no solo la vacuna en sí, en condiciones naturales de campo en lugar de en un ensayo clínico controlado. La eficacia de la vacuna es proporcional a la eficacia de la vacuna (potencia), pero también se ve afectada por la eficacia de la inmunización de los grupos objetivo de la población, así como por otros factores no relacionados con la vacuna que influyen en los resultados del 'mundo real' de las hospitalizaciones y las visitas ambulatorias, o los costos. Por ejemplo, se puede utilizar un análisis retrospectivo de casos y controles, en el que se comparan las tasas de vacunación entre un conjunto de casos infectados y los controles apropiados. La eficacia de la vacuna se puede expresar como una diferencia de tasas, con el uso de la razón de posibilidades (OR) para desarrollar una infección a pesar de la vacunación:

$$\text{Efectividad} = (1 - \text{OR}) \times 100.$$

En algunas modalidades, la eficacia de la composición inmunizante (p. ej., vacuna de ARN) es de al menos el 60 % en relación con los sujetos de control no vacunados. Por ejemplo, la eficacia de la composición inmunizante puede ser al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 95 %, al menos 98 % o 100 % en relación con sujetos de control no vacunados.

Inmunidad esterilizante. La inmunidad esterilizante se refiere a un estado inmunológico único que previene la infección efectiva por patógenos en el huésped. En algunas modalidades, la cantidad eficaz de una composición inmunizante de la presente descripción es suficiente para proporcionar inmunidad esterilizante en el sujeto durante al menos 1 año. Por ejemplo, la cantidad eficaz de una composición inmunizante de la presente descripción es suficiente para proporcionar inmunidad esterilizante en el sujeto durante al menos 2 años, al menos 3 años, al menos 4 años o al menos 5 años. En algunas modalidades, la cantidad eficaz de una composición inmunizante de la presente descripción es suficiente para proporcionar inmunidad esterilizante en el sujeto a una dosis al menos 5 veces menor con respecto al control. Por ejemplo, la cantidad eficaz puede ser suficiente para proporcionar inmunidad esterilizante en el sujeto a una dosis al menos 10 veces menor, 15 veces menor o 20 veces menor con respecto a un control.

Antígeno detectable. En algunas modalidades, la cantidad eficaz de una composición

inmunizante de la presente descripción es suficiente para producir niveles detectables de antígeno de coronavirus medidos en el suero del sujeto 1-72 horas luego de la administración.

Título. Un título de anticuerpos es una medida de la cantidad de anticuerpos dentro de un sujeto, por ejemplo, anticuerpos que son específicos para un antígeno particular (p. ej., un antígeno anticoronavirus). El título de anticuerpos se expresa típicamente como la inversa de la mayor dilución que proporciona un resultado positivo. El ensayo inmunoabsorbente ligado a enzimas (ELISA) es un ensayo común para determinar los títulos de anticuerpos, por ejemplo.

En algunas modalidades, la cantidad eficaz de una composición inmunizante de la presente descripción es suficiente para producir un título de anticuerpo neutralizante 1000-10 000 producido por anticuerpo neutralizante contra el antígeno del coronavirus medido en el suero del sujeto 1-72 horas luego de la administración. En algunas modalidades, la cantidad eficaz es suficiente para producir un título de anticuerpo neutralizante de 1000-5000 producido por anticuerpo neutralizante contra el antígeno de coronavirus medido en el suero del sujeto 1-72 horas luego de la administración. En algunas modalidades, la cantidad eficaz es suficiente para producir un título de anticuerpo neutralizante de 5000-10 000 producido por anticuerpo neutralizante contra el antígeno del coronavirus medido en el suero del sujeto entre 1 y 72 horas luego de la administración.

En algunas modalidades, el título de anticuerpos neutralizantes es de al menos 100 NT₅₀. Por ejemplo, el título de anticuerpos neutralizantes puede ser de al menos 200, 300, 400, 500, 600, 700, 800, 900 o 1000 NT₅₀. En algunas modalidades, el título de anticuerpos neutralizantes es de al menos 10 000 NT₅₀.

En algunas modalidades, el título de anticuerpos neutralizantes es de al menos 100 unidades neutralizantes por mililitro (NU/ml). Por ejemplo, el título de anticuerpo neutralizante puede ser de al menos 200, 300, 400, 500, 600, 700, 800, 900 o 1000 NU/ml. En algunas modalidades, el título de anticuerpos neutralizantes es de al menos 10 000 NU/ml.

En algunas modalidades, un título de anticuerpo antigénico anticoronavirus producido en el sujeto aumenta en al menos 1 log con respecto a un control. Por ejemplo, un título de anticuerpo antigénico anticoronavirus producido en el sujeto puede incrementarse en al menos 2, 3, 4, 5, 6, 7, 8, 9 o 10 log con respecto a un control.

En algunas modalidades, un título de anticuerpo de antígeno anticoronavirus producido en el sujeto aumenta al menos 2 veces con respecto a un control. Por ejemplo, un título de anticuerpo antigénico anticoronavirus producido en el sujeto aumenta al menos 3, 4, 5, 6, 7, 8, 9 o 10 veces con respecto a un control.

- 5 En algunas modalidades, una media geométrica, que es la raíz enésima del producto de n números, se utiliza generalmente para describir el crecimiento proporcional. La media geométrica, en algunas modalidades, se utiliza para caracterizar el título de anticuerpo producido en un sujeto.

- 10 Un control puede ser, por ejemplo, un sujeto no vacunado o un sujeto al que se le haya administrado una vacuna vírica viva atenuada, una vacuna vírica inactivada o una vacuna de subunidad proteica.

EJEMPLOS

15

Ejemplo 1: Expresión *in vitro* de nCoV - ADN

Las construcciones probadas en este experimento fueron el ADN de Norwood cotransfectado con un plásmido de polimerasa T7 para transactivar el promotor en el plásmido 2019-nCoV de Norwood. Se utilizó un ADN de control positivo para el SARS. Las

- 20 condiciones del ensayo fueron las siguientes:

Construcciones de ADN:	SARS-CoV-2 Variantes 6-10
Tipo de célula:	Células HEK293T
Formato de placa:	12 pocillos a 600 000 células/pocillo
ADN por pocillo:	2.5 µg/pocillo (construcción: T7 = 1:1)
25 Tiempo de incubación:	24, 72 horas
Tinción extracelular:	de un solo color
Instrumento:	LSR Fortessa
ACE2-FLAG, His:	200 µg de stock, 10 µg/ml de concentración de FACS
Anti FLAG-FITC:	1 mg, 5 µg/ml de concentración de FACS

30

Ejemplo 2: Expresión *in vitro* de nCoV - ARNm

Se ensayó el ARNm de las construcciones del Ejemplo 1. Las condiciones del ensayo fueron

las siguientes:

	Construcciones de ARNm:	Variantes 6-10 de SARS-CoV-2
	Tipo de célula:	Células HEK293T
	Formato de la placa:	24 pocillos a 300 000 células/pocillo
5	ARNm por pocillo:	0.5 µg, 0.1 µg/pocillo
	Tiempo de incubación:	24, 48 horas
	Tinción extracelular:	de un solo color
	Instrumento:	LSR Fortessa
	ACE2-FLAG, His:	200 µg de stock, 10 µg/ml de concentración de FACS
10	Anti FLAG-FITC:	1 mg, 5 µg/ml de concentración de FACS

Entre todas las construcciones, la variante 5 de SARS-CoV-2 mostró la mejor expresión en comparación con otras en dosis bajas. Véanse las Figuras 2 y 3.

Ejemplo 3: Estudio de inmunogenicidad

15 El presente estudio está diseñado para probar la inmunogenicidad en ratones y/o conejos de las vacunas de coronavirus candidatas que comprenden un ARNm de la **Tabla 1** que codifica un antígeno de coronavirus (p. ej., la proteína pico (S), la subunidad S1 (S1) de la proteína de pico, o la subunidad S2 (S2) de la proteína de pico), como un antígeno del SARS-CoV-2.

20 Los animales se vacunan en las semanas 0 y 3 por vía intravenosa (IV), intramuscular (IM) o intradérmica (ID). Como controles, un grupo permanece sin vacunar y a otro se le administra coronavirus inactivado. Se recolecta suero de cada animal en las semanas 1, 3 (antes de la dosis) y 5. Se analizan los sangrados individuales para determinar la actividad anti S, anti S1 o anti S2 mediante un ensayo de neutralización de virus de los tres puntos
25 de tiempo y muestras combinadas a partir de la semana 5 solo se prueban mediante Western blot utilizando coronavirus inactivado.

En experimentos en los que se utiliza una formulación de nanopartículas lipídicas (LNP), la formulación puede incluir lípidos modificados con PEG al 0.5-15 %; 5-25 % de lípidos no catiónicos; 25-55 % de esterol; y 20-60 % de lípidos catiónicos ionizables. El lípido
30 modificado con PEG puede ser 1,2 dimiristoil-sn-glicerol, metoxipolietilenglicol (PEG2000 DMG), el lípido no catiónico puede ser 1,2 diestearoil-sn-glicero-3-fosfocolina (DSPC), el esterol puede ser colesterol; y el lípido catiónico ionizable puede tener la estructura del

Compuesto 1, por ejemplo.

Ejemplo 4: Exposición al coronavirus

El presente estudio está diseñado para probar la eficacia en ratones y/o conejos de vacunas de coronavirus candidatas que comprenden un ARNm de la **Tabla 1** que codifica un antígeno de coronavirus (p. ej., la proteína de pico (S), la subunidad S1 (S1) de la proteína de pico, o la subunidad S2 (S2) de la proteína de pico), como un antígeno del SARS-CoV-2, contra una exposición letal con un coronavirus. Los animales se expusieron a una dosis letal (10xLD90; ~ 100 unidades formadoras de placa; PFU) de coronavirus.

Los animales utilizados son hembras de 6-8 semanas de edad en grupos de 10. Los animales se vacunan en las semanas 0 y 3 mediante una vía de administración IM, ID o IV. Las vacunas candidatas están químicamente modificadas o sin modificar. El suero animal se prueba para la microneutralización (véase el Ejemplo 14). A continuación, los animales se expusieron a ~ 1 LD90 de coronavirus en la semana 7 mediante una vía de administración IN, IM, ID o IV. El criterio de valoración es el día 13 luego de la infección, muerte o eutanasia. Los animales que presentan una enfermedad grave determinada por una pérdida de peso > 30 %, letargo extremo o parálisis se sacrifican. La temperatura corporal y el peso se evalúan y registran diariamente.

Ejemplo 5 - Inmunogenicidad de la Variante 9 del SARS-CoV-2 en ratones (una dosis)

Se inmunizaron ratones C3B6, C57/BL6 y BALB/c con varias dosis de la vacuna de ARNm de la variante 9 del SARS-CoV-2 ("Variante 9") en 50 µL de 1X PBS inyectados por vía intramuscular en el lado derecho pata trasera. Dos semanas luego de la inmunización, se recogieron sueros y se sometieron a ELISA para evaluar la unión del anticuerpo a la proteína de pico de prefusión estabilizada con SARS-CoV-2 (SARS-CoV-2 pre-S).

Los datos de la Variante 9 se muestran en las Figuras 4A-4B. No hubo diferencias significativas entre las cepas de ratón. Como se muestra en la Figura 4A, los ratones C3B6 que recibieron 1 µg de Variante 9 tuvieron respuestas de anticuerpos significativamente más altas que los ratones C3B6 que recibieron dosis de 0.1 µg o 0.01 µg (valor de p <0.05). La Figura 4B muestra que los ratones BALB/c que recibieron 10 µg de Variante 9 tuvieron respuestas de anticuerpos significativamente más altas que los ratones BALB/c que recibieron la dosis de 1 µg (valor p <0.05) o la dosis de 0.1 µg (valor p <0.0001).

Otros candidatos de ARNm se probaron de la manera descrita anteriormente, como se muestra en las Figuras 5A-5C. La Figura 5A demuestra que la vacuna de ARNm de la variante 5 del SARS-CoV-2 ("Variante 5") provocó respuestas de anticuerpos similares en ratones C3B6 y BALB/c luego de la administración de una dosis. Los ratones BALB/c que recibieron 1 µg de Variante 9 o Variante 5 tuvieron respuestas de anticuerpos significativamente más altas en comparación con los ratones BALB/c que recibieron dosis de 0.1 µg o 0.01 µg (valor $p < 0.05$) (Figura 5B). A 0.1 µg, la Variante 9 provocó respuestas similares a varios otros antígenos de la vacuna del SARS-CoV-2 administrados por ARNm. Además, a la dosis de 0.1 µg, la vacuna de ARNm del SARS-CoV-2 variante 8 ("Variante 8") y la vacuna de ARNm del SARS-CoV-2 variante 6 ("Variante 6") provocaron respuestas de anticuerpos significativamente más altas que la proteína soluble de pico (S) secuencias (* = valor $p < 0.05$; ** = valor $p < 0.01$).

Análisis temporal

Se inmunizaron ratones BALB/c con varias dosis de la Variante 9 fabricada mediante el proceso clínicamente representativo en 50 µL de 1X PBS por vía intramuscular en la pata trasera derecha. Cada dos semanas, se recolectaron sueros postcebo y se sometieron a ELISA para evaluar la unión del anticuerpo a la proteína de pico de prefusión estabilizada con SARS-CoV-2 (SARS-CoV-2 pre-S). Los resultados se muestran en la Figura 6. Cada símbolo representa el título medio geométrico (GMT) y las barras de error indican la desviación estándar geométrica (SD). Se utilizó ANOVA bidireccional para comparar las respuestas de anticuerpos a lo largo del tiempo y cada dosis. Se encontró que la dosis de 10 µg de Variante 9 provocaba una respuesta de anticuerpos significativamente más alta que todas las demás dosis (valor $p < 0.0001$) y decae significativamente durante 4 semanas luego del cebado (valor $p < 0.001$).

Ejemplo 6: inmunogenicidad de la variante 9 en ratones (dos dosis)

Comparación de cepas de ratón

Se inmunizaron ratones (BALB/c, C57BL/6 y C3B6) con varias dosis de Variante 9 en 50 µL de 1X PBS por vía intramuscular en la pata trasera derecha en las semanas 0 y 3 (Figura 7). Dos semanas luego del cebado y del refuerzo (p. ej., semanas 2 y 5), se recogieron sueros y se sometieron a ELISA para evaluar la unión del anticuerpo a la proteína de pico de

prefusión estabilizada con SARS-CoV-2 (SARS-CoV-2 pre-S).

Los resultados se muestran en las Figuras 8A-8C. Cada símbolo representa un ratón individual, las barras representan títulos de media geométrica (GMT) y las barras de error indican la desviación estándar geométrica (SD). Se utilizó ANOVA bidireccional para
5 comparar las respuestas postcebo y posrefuerzo. A la dosis de 1 µg, los ratones BALB/c (Figura 8A) y C57/BL6 (Figura 8B) inmunizados con la Variante 9 tuvieron respuestas de anticuerpos significativamente más altas luego del refuerzo (valor $p < 0.0001$).

Comparación de las construcciones de la vacuna de ARNm del SARS-CoV-2

Se inmunizaron ratones (BALB/c y C3B6) con varias dosis de la Variante 9, Variante 5 o
10 ARNm de proteína de pico de tipo salvaje del SARS-CoV-2 en 50 µl de 1X PBS por vía intramuscular en la pata trasera derecha en las semanas 0 y 3 (Figura 7). Dos semanas luego del cebo y del refuerzo (p. ej., semanas 2 y 5), se recogieron sueros y se sometieron a ELISA para evaluar la unión del anticuerpo a la proteína de pico de prefusión estabilizada con SARS-CoV-2 (SARS-CoV-2 pre-S).

Los resultados se muestran en las Figuras 9A-9E. Cada símbolo representa un ratón individual, las barras representan títulos de media geométrica (GMT) y las barras de error indican la desviación estándar geométrica (SD). Se utilizó ANOVA bidireccional para
15 comparar las respuestas postcebo y posrefuerzo. A la dosis de 1 µg, los ratones inmunizados con la Variante 5 y la secuencia de tipo salvaje de pico (S WT) tuvieron respuestas de anticuerpos significativamente más altas luego del refuerzo (valor $p < 0.0001$) (Figuras 9A-9C). Además, a la dosis de 1 µg, los ratones BALB/c inmunizados con la Variante
20 9 tuvieron respuestas de anticuerpos significativamente más altas que los ratones inmunizados con la secuencia de respaldo GMP (valor $p < 0.01$) y el S WT (valor $p < 0.05$) (Figura 9D). No hubo diferencia significativa entre las respuestas de anticuerpos provocadas
25 por cualquiera de las construcciones en los ratones C3B6 (Figura 9E). Para las tres secuencias probadas, hubo una respuesta temporal significativa (p. ej., dosis postcebo frente a dosis postrefuerzo).

Se ensayaron otras secuencias de investigación. Se inmunizaron ratones BALB/c con varias dosis de ARNm que codifica la Variante 9 u otras secuencias de investigación en 50 µL de
30 1X PBS por vía intramuscular en la pata trasera derecha en las semanas 0 y 3 (Figura 7). Dos semanas luego del refuerzo (p. ej., semana 5), se recogieron sueros y se sometieron a ELISA para evaluar la unión del anticuerpo a la proteína de pico de prefusión estabilizada

con SARS-CoV-2 (SARS-CoV-2 pre-S).

Los resultados se muestran en la Figura 10. Cada símbolo representa un ratón individual, las barras representan títulos de media geométrica (GMT) y las barras de error indican la desviación estándar geométrica (SD). Se utilizó un ANOVA de una vía para comparar todos los inmunógenos. Los ratones inmunizados con la Variante 8, la Variante 7 y la Variante 6 tenían títulos de anticuerpos significativamente más altos que los ratones inmunizados con la variante 9 y S WT (* = valor $p < 0.05$; ** = valor $p < 0.01$; **** = p -valor < 0.0001).

Ejemplo 7 - Expresión *in vivo* de construcciones de vacuna de ARNm de SARS-CoV-2

A ratones hembra BALB/c, de 6 a 8 semanas de edad, se les administraron 2 μ g o 10 μ g de una construcción COVID-19 o amortiguador Tris (como control) por vía intramuscular en cada pata trasera. Las construcciones comprenden la variante 9 en nanopartículas de lípidos catiónicos, acetato de sodio 10.7 mM, sacarosa al 8.7 %, Tris 20 mM (pH 7.5). Se probaron tres construcciones: Variante 9, Variante 5 y Variante 6. Las construcciones se almacenaron a -70 °C (Variante 9) o -20 °C (otras construcciones). Un día después, se recogieron bazo y ganglios linfáticos para detectar la expresión de proteínas mediante citometría de flujo.

Las Figuras 11A-11B muestran los resultados utilizando el anticuerpo 5653-118 ("118"), que es específico para el dominio N-terminal de la subunidad S1 de SARS-CoV-1. Hubo una buena expresión de todas las construcciones probadas y se observó una caída en la expresión dependiente de la dosis. En el ganglio linfático (Figura 11A) y en el bazo (Figura 11B), la Variante 5 mostró una expresión significativamente más alta ($\alpha = 0.05$) que cualquiera de las otras construcciones. A la dosis más baja (2 μ g), la Variante 9 tuvo una expresión significativamente mayor ($\alpha = 0.05$) que la Variante 6.

Las Figuras 12A-12B muestran los resultados utilizando el anticuerpo 5652-109 ("109"), que es específico para el dominio de unión a la proteína S de SARS-CoV-1. Hubo una buena expresión de todas las construcciones probadas y se observó una caída en la expresión dependiente de la dosis. No hubo diferencia significativa entre la Variante 9 y la Variante 5 a la dosis de 10 μ g en el ganglio linfático o el bazo. A la dosis de 2 μ g, la Variante 9 tenía una expresión significativamente mayor ($\alpha = 0.05$) que las otras dos construcciones en el ganglio linfático (Figura 12A) y en el bazo (Figura 12B).

Ejemplo 8 - Expresión *in vitro* de construcciones de vacuna de ARNm de SARS-CoV-2

Se probaron seis construcciones de vacunas de ARNm del SARS-CoV-2 *in vitro*. Se sembraron células HEK293t (30 000 células/pocillo) en una placa de 96 pocillos. Se añadieron 200 ng de la construcción a cada pocillo y las placas se incubaron durante 24 horas. Luego, las células se tiñeron con el anticuerpo "118" (en diluciones de 1:100, 1:300 o 1:600), el anticuerpo "109" (en diluciones de 1:100, 1:300 o 1:600), o SARS-103 (se une a RBD a partir de SARS-CoV-1 a una dilución de 1:100; como control). A continuación, se realizó la tinción con FC AL647 antihumano, a una dilución de 1:500, y se leyeron las muestras utilizando LSR Fortessa. Los resultados se muestran en las Figuras 13A-13C. Los resultados de la variante 9 se proporcionan en la Figura 14

10

Ejemplo 9 - desarrollo de un ensayo de potencia *in vitro*

Se desarrolló un ensayo para examinar la potencia de diferentes construcciones. Se probaron dos anticuerpos específicos, 118 (específicos para el dominio N-terminal de la subunidad S1 de SARS-CoV-1) y 109 (específicos para el dominio de unión al receptor de la proteína S de SARS-CoV-1). Como se muestra en la Figura 15, solo el anticuerpo 118 se unió al antígeno del SARS-CoV-2 a diferentes concentraciones y dosis.

15

Ejemplo 10 - Estudio de inmunogenicidad en ratón con vacuna de ARNm de la variante 9 del SARS-CoV-2

Se iniciaron estudios para evaluar la inmunogenicidad y la eficacia de dosis bajas de la vacuna de ARNm de la variante 9 del SARS-CoV-2 ("Variante 9") en ratones BALB/c. Se vacunaron ratones BALB/c con 1 µg, 0.1 µg o 0.01 µg de Variante 9 en las semanas 0 y 3. Los anticuerpos de unión a la proteína S-2P estabilizada se cuantificaron en las semanas 2 y 5. Dos semanas después de una dosis única, hubo niveles sustanciales de anticuerpos de unión a la proteína S-2P medidos por ELISA en ratones que recibieron 1 µg de Variante 9 (Figura 16A). Una segunda dosis de Variante 9 aumentó significativamente el nivel de anticuerpos de unión en ratones que recibieron 1 µg o 0.1 µg de Variante 9 (Figura 16A). La actividad neutralizante contra el SARS-CoV-2 se evaluó utilizando un ensayo de neutralización del indicador de lentivirus pseudotipado. La Variante 9 provocó una actividad neutralizante significativa a la dosis de 1 µg en comparación con los ratones que recibieron 0.1 µg de la Variante 9 (Figura 16B).

25

30

Se inmunizaron ratones BALB/c con 1 o 0.1 µg de Variante 9, Variante 5 o de tipo salvaje

(WT) sin la mutación de prolina 2 en las semanas 0 y 3. Dos semanas después del refuerzo, los sueros se recogieron y analizaron mediante ensayos de neutralización de lentivirus reporteros seudotipados contra el SARS-CoV-2 homólogo. Se generaron curvas sigmoideas, tomando promedios de triplicados en cada dilución de suero, a partir de lecturas de unidades relativas de luciferasa (RLU) y Se calculó la actividad neutralizante de 21B, 21D, 21F, 21H) al 50 % (IC₅₀) (Figuras 21A, 21C, 21E, 21G) y 80 % (IC₈₀) (Figuras 21B, 21D, 21F, 21H) considerando que las células no infectadas representan el 100 % de neutralización y las células transducidas con solo virus para representar el 0 % de neutralización. Como se muestra en las Figuras 21A-21F, los ratones inmunizados con 1 µg de ARNm de SARS-CoV-2 S tuvieron respuestas de anticuerpos neutralizantes significativamente más altas que los ratones inmunizados con 0.1 µg (** = valor p <0.001, *** = valor p <0.0001). Además, los ratones inmunizados con una dosis de 0.1 µg no tenían actividad neutralizante detectable, lo que sugiere niveles de anticuerpos subprotectores. Además, el SARS-CoV-2 S estabilizado con mutación 2P (Variante 5 y Variante 9) indujo una actividad neutralizante de IC₅₀ más potente que WT S (* valor p <0.05). La inclusión del sitio de escisión de furina S1/S2 nativo o la sustitución del sitio de escisión de furina con un enlazador GS para producir una construcción de cadena sencilla no tuvo un efecto significativo sobre la inmunogenicidad (Figuras 21G, 21H).

Se expusieron ratones BALB/c inmunizados con 1 µg, 0.1 µg, o 0.01 µg de la Variante 9 en las semanas 0 y 3 en 50 µl de 1X PBS por vía intramuscular en la pata trasera derecha en la semana 9 por vía intranasal con 1 x 10⁵ PFU de SARS-CoV-2 adaptado de un ratón que contiene dos cambios de aminoácidos dirigidos en el dominio de unión al receptor para eliminar los choques con el receptor ACE-2 de ratón (véase la Figura 19 para el programa). El día 2 después de la exposición, se homogeneizaron los pulmones y la nariz de los ratones y se evaluó la carga viral mediante ensayos de placa. Las unidades formadoras de placa en un lóbulo de pulmón (Figura 17A) y en cornetes nasales (Figura 17B) muestran que el grupo de dosis de 1 µg está completamente protegido con una reducción de 60 veces en el título en comparación con el grupo de control. Por el contrario, los ratones expuestos no inmunizados tenían cargas virales de aproximadamente 10⁶ PFU por lóbulo pulmonar. Se observó un efecto de dosis: la dosis de 0.1 µg de la Variante 9 redujo la carga viral pulmonar en aproximadamente 2 logaritmos y la dosis de 0.01 µg de la Variante 9 redujo la carga viral pulmonar en aproximadamente 0.5 log.

En un estudio adicional, los ratones fueron vacunados con 10 µg, µg o 0.1 µg de Variante 9 en 50 µL de 1X PBS por vía intramuscular en la pata trasera derecha una vez (semana 0) y se expusieron intranasalmente en la semana 7 con 1×10^5 PFU de un SARS-CoV-2 adaptado al ratón que contiene dos cambios de aminoácidos dirigidos en el dominio de unión al receptor para eliminar los choques con el receptor ACE-2 del ratón. El día 2 después de la exposición, se homogeneizaron los pulmones de ratón (Figura 18A) y la nariz (Figura 18B) y se evaluó la carga viral mediante un ensayo de placa. Como se observa en la Figura 18A, los grupos de dosis de 10 µg y de dosis de 1 µg están completamente protegidos de la replicación viral en el pulmón luego de la exposición, con una reducción de 60 veces en el título en comparación con el grupo de control. El porcentaje de peso corporal se muestra en la Figura 18C.

En un estudio adicional, los ratones se vacunaron con 1 µg, 0.1 µg o 0.01 µg de Variante 9 en las semanas 0 y 4 y se expusieron en la semana 7 con un SARS-CoV-2 adaptado al ratón que contiene dos cambios de aminoácidos dirigidos en el dominio de unión al receptor para eliminar los choques con el receptor ACE-2 de ratón. Las unidades formadoras de placa en un lóbulo del pulmón (Figura 20A) y en los cornetes nasales (Figura 20B) y en el día dos después de la exposición muestran que el grupo de dosis de 1 µg y el grupo de dosis de 0.1 µg están completamente protegidos, con aproximadamente una Reducción de 60 veces en el título en comparación con el grupo de control. El porcentaje de peso corporal se muestra en la Figura 20C.

Ejemplo 11 - Vacuna de ARNm de la Variante 9 del SARS-CoV-2 en comparación con secuencias alternativas

Este Ejemplo proporciona datos relacionados con la unión y neutralización de las respuestas de anticuerpos luego de la inmunización de ARNm de dosis baja con diseños de antígenos de punta alternativos. Se inmunizaron ratones BALB/c con 0.1 µg de ARNm que codifica diferentes variantes de SARS-CoV-2 S-2P. Los ratones se inmunizaron dos veces en las semanas 0 y 3. Dos semanas luego del refuerzo, se recogieron los sueros y se analizaron mediante ELISA competitivo de doblez contra la punta estabilizada con SARS-CoV-2 homóloga (Figura 22A) y ensayos de neutralización del indicador de lentivirus pseudotipado (Figura 22B). La Figura 22A muestra los títulos de unión del punto final del suero, obtenidos tomando promedios de duplicados de cada dilución de suero, y calculados como 4 veces por

encima de la densidad óptica de fondo. Además, las curvas sigmoideas, teniendo las medias de triplicados en cada dilución de suero, se generaron a partir de unidades relativas de luciferasa (RLU) lecturas un 50 % (IC_{50}) se calculó actividad neutralizante teniendo en cuenta las células no infectadas que representan 100 % de neutralización y las células transducidas con sólo el virus que representa 0 % de neutralización (Figura 22B). Además, los títulos de neutralización y unión de anticuerpos se compararon mediante correlación de Spearman (Figura 22C). Se encontró que las secuencias que codifican el ARNm que contienen mutaciones de la cola citoplasmática provocan las respuestas de anticuerpos más potentes. Además, hubo una fuerte correlación entre los títulos de anticuerpos de unión y los títulos de anticuerpos neutralizantes, cuando correspondía.

Métodos

ELISA de SARS-CoV-2

Para medir la unión de anticuerpos, se realizó un ELISA. Se recubrió SARS-CoV-2 pre-S sobre placas de fondo plano Nunc MaxiSorp™ de 96 pocillos (ThermoFisher, núm. de catálogo: 44-2401-21) en 100 μ L de 1X PBS durante 16 horas a 4 °C. Las placas se lavaron 3 veces con

250 μ L de PBS-Tween (PBST) (Medicago AB, núm. de catálogo: 09-9410-100). Para evitar la unión inespecífica, las placas se bloquearon con 200 μ L de PBST suplementado con leche descremada al 5 % (BD Difco™, núm. de catálogo: 232100) (amortiguador de bloqueo) durante 1 hora a temperatura ambiente (TA). Las placas se lavaron 3 veces con 250 μ L de PBST. Los sueros se diluyeron en serie (1:100, 4 veces, 8 veces) en 100 μ L de amortiguador de bloqueo y se dejó que se unieran al antígeno durante 1 hora a TA, por duplicado. Las placas se lavaron 3 veces con 250 μ L de PBST. Se añadieron 100 μ L de anticuerpo secundario de adsorción cruzada de IgG de cabra (H+L) antirratón conjugado con HRP (ThermoFisher, núm. de catálogo: G-21040) diluido en amortiguador de bloqueo durante 1 hora a TA. Las placas se lavaron 3 veces con 250 μ L de PBST. La reacción ligada a enzimas se desarrolló durante 10 minutos con 100 μ L de sustrato de peroxidasa de micropocillos TMB de 1 componente KPL SureBlue (Sure Blue, núm. de catálogo: 5120-0077) y se detuvo con 100 μ L de ácido sulfúrico 1N (ThermoFisher, núm. de catálogo: SA 212-1). Se utilizó el paradigma Spectramax (dispositivos moleculares) para detectar OD₄₅₀. Los títulos de criterio de valoración del suero se calcularon como 4 veces por encima de la unión del anticuerpo

secundario no específico al antígeno.

Modalidades adicionales

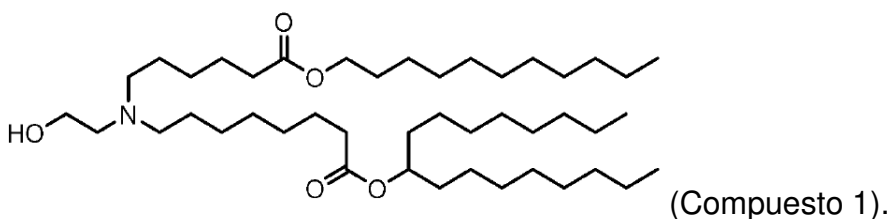
1. Un ácido ribonucleico (ARN) que comprende un marco de lectura abierto (ORF) que
5 codifica un antígeno de coronavirus capaz de inducir una respuesta inmune, como una
respuesta de anticuerpos neutralizantes, al SARS-CoV-2, opcionalmente en el que el ARN
se formula en un nanopartícula lipídica.
2. Un ácido ribonucleico (ARN) químicamente modificado que comprende un marco de
lectura abierto (ORF) que comprende una secuencia que tiene al menos un 80 % de
10 identidad con un ARN de tipo salvaje que codifica un antígeno del SARS-CoV-2, en donde
opcionalmente se formula el ARN en una nanopartícula lipídica.
3. Un ácido ribonucleico (ARN) de codón optimizado que comprende un marco de
lectura abierto (ORF) que comprende una secuencia que tiene al menos un 80 % de
identidad con un ARN de tipo salvaje que codifica un antígeno del SARS-CoV-2, en donde
15 opcionalmente se formula el ARN en una nanopartícula lipídica.
4. El ARN del párrafo 2 o 3, en donde el antígeno del SARS-CoV-2 codificado por el
ARN de tipo salvaje comprende la secuencia de SEQ ID NO: 31.
5. Un ácido ribonucleico (ARN) que comprende un marco de lectura abierto (ORF) que
comprende una secuencia que tiene al menos 80 % de identidad con la secuencia de
20 cualquiera de las SEQ ID NO: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62,
64, 66, 68, 70, 72, 74, 76, 78, 80, 82 u 84.
6. El ARN del párrafo 5, en donde el ORF comprende una secuencia que tiene al menos
el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la
secuencia de cualquiera de las SEQ ID NO: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52,
25 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 u 84.
7. El ARN del párrafo 5 o 6 comprende además una UTR 5', en donde opcionalmente
la UTR 5' comprende la secuencia de SEQ ID NO: 2 o SEQ ID NO: 36.
8. El ARN de cualquiera de los párrafos precedentes que comprende además una UTR
3', en donde opcionalmente la UTR 3' comprende la secuencia de SEQ ID NO: 4 o SEQ ID
30 NO: 37.
9. El ARN de cualquiera de los párrafos precedentes que comprende además un
análogo de cubierta 5', opcionalmente una cubierta de 7mG (5') ppp (5') NImpNp.

10. El ARN de cualquiera de los párrafos precedentes que comprende además una cola de poli(A), que opcionalmente tiene una longitud de 50 a 150 nucleótidos.
11. El ARN de cualquiera de los párrafos 5 a 10, en donde el ORF codifica un antígeno del SARS-CoV-2.
- 5 12. El ARN del párrafo 11, en donde el antígeno del coronavirus es una proteína estructural.
13. El ARN del párrafo 12, en donde la proteína estructural es una proteína de pico.
14. El ARN de cualquiera de los párrafos 11-13, en donde el antígeno del coronavirus comprende una secuencia que tiene al menos 80 % de identidad con la secuencia de
10 cualquiera de las SEQ ID NO: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83 u 85.
15. El ARN del párrafo 14, en donde el antígeno del coronavirus comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de cualquiera de las SEQ ID NO: 5, 8, 11, 14, 17, 20,
15 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83 u 85.
16. El ARN de cualquiera de los párrafos 1-13, en donde el ORF comprende la secuencia de cualquiera de las SEQ ID NO: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 u 84.
17. El ARN de cualquiera de los párrafos 1-13, en donde el ARN comprende una
20 secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de cualquiera de las SEQ ID NO: 1, 6, 9, 12, 15, 18, 21, 24, 27, 30, 51, 53, 55, 57-58, 60 u 86-97.
18. El ARN de cualquiera de los párrafos 1-13, en donde el RNA comprende la secuencia de cualquiera de las SEQ ID NO: 1, 6, 9, 12, 15, 18, 21, 24, 27, 30, 51, 53, 55, 57-58, 60 u
25 86-97.
19. El ARN de cualquiera de los párrafos precedentes, en donde el ARN comprende una modificación química y opcionalmente está completamente modificado químicamente.
20. El ARN del párrafo 19, en donde la modificación química es 1-metilpseudouridina y opcionalmente cada uridina es 1-metilpseudouridina.
- 30 21. El ARN del párrafo 19, en donde cada uridina es 1-metilpseudouridina.
22. El ARN de cualquiera de los párrafos precedentes formulado en una nanopartícula lipídica.

23. El ARN del párrafo 22, en donde la nanopartícula lipídica comprende un lípido modificado con PEG, un lípido no catiónico, un estero, un lípido catiónico ionizable o cualquier combinación de estos.

24. El ARN del párrafo 23, en donde la nanopartícula lipídica comprende 0.5-15 % en moles de lípido modificado con PEG; 5-25 % en moles de lípidos no catiónicos; 25-55 % en moles de estero; y 20-60 % en moles de lípido catiónico ionizable.

25. El ARN del párrafo 24, en donde el lípido modificado con PEG es 1,2 dimiristoil-sn-glicerol, metoxipolietilenglicol (PEG2000 DMG), el lípido no catiónico es 1,2 diestearoil-sn-glicero-3-fosfocolina (DSPC), el estero es colesterol; y el lípido catiónico ionizable tiene la estructura del Compuesto 1:

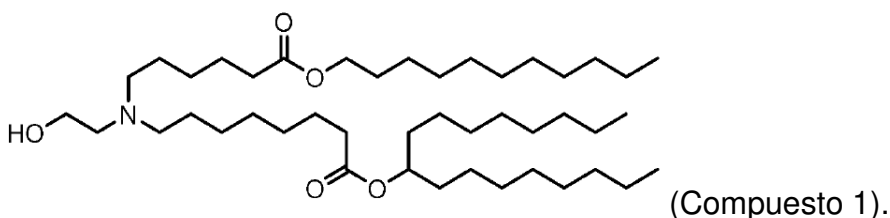


26. Una composición que comprende el ARN de cualquiera de los párrafos 1-21 y una mezcla de lípidos.

27. La composición del párrafo 26, en donde la mezcla de lípidos comprende un lípido modificado con PEG, un lípido no catiónico, un estero, un lípido catiónico ionizable o cualquier combinación de estos.

28. La composición del párrafo 27, en donde la mezcla de lípidos comprende 0.5-15 % en moles de lípido modificado con PEG; 5-25 % en moles de lípidos no catiónicos; 25-55 % en moles de estero; y 20-60 % en moles de lípido catiónico ionizable.

29. La composición del párrafo 28, en donde el lípido modificado con PEG es 1,2 dimiristoil-sn-glicerol, metoxipolietilenglicol (PEG2000 DMG), el lípido no catiónico es 1,2 diestearoil-sn-glicero-3-fosfocolina (DSPC), el estero es colesterol; y el lípido catiónico ionizable tiene la estructura del Compuesto 1:



30. La composición de cualquiera de los párrafos 26-29, en donde la mezcla de lípidos

forma nanopartículas lipídicas.

31. La composición del párrafo 30, en donde el ARN se formula en las nanopartículas lipídicas.

32. El ARN de cualquiera de los párrafos 1-13, en donde el ORF comprende una
5 secuencia de nucleótidos que tiene al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 98 % o 100 % de identidad con la secuencia de SEQ ID NO: 28.

33. El ARN de cualquiera de los párrafos 1-13, en donde el antígeno del coronavirus
comprende una secuencia de aminoácidos que tiene al menos 80 %, al menos 85 %, al
menos 90 %, al menos 95 %, al menos 98 % o 100 % de identidad con la secuencia de SEQ
10 ID NO: 29.

33. El ARN de cualquiera de los párrafos 1-13, en donde el RNA comprende una
secuencia de nucleótidos que tiene al menos 85 %, al menos 90 %, al menos 95 %, al menos
98 % o 100 % de identidad con la secuencia de SEQ ID NO: 27.

34. Un método que comprende administrar a un sujeto el ARN o composición de
15 cualquiera de los párrafos precedentes en una cantidad eficaz para inducir una respuesta
de anticuerpo neutralizante contra un coronavirus en el sujeto.

35. Un método que comprende administrar a un sujeto el ARN o composición de
cualquiera de los párrafos precedentes en una cantidad eficaz para inducir una respuesta
de anticuerpos neutralizantes y/o una respuesta inmune de células T, opcionalmente una
20 respuesta inmune de células T CD4⁺ y/o CD8⁺ frente a un coronavirus en el sujeto.

36. El método de los párrafos 34 y 35, en donde el coronavirus es un SARS-CoV-2.

37. El método de cualquiera de los párrafos de método precedentes, en donde el sujeto
está inmunodeprimido.

38. El método de cualquiera de los párrafos de método precedentes, en donde el sujeto
25 tiene una enfermedad pulmonar.

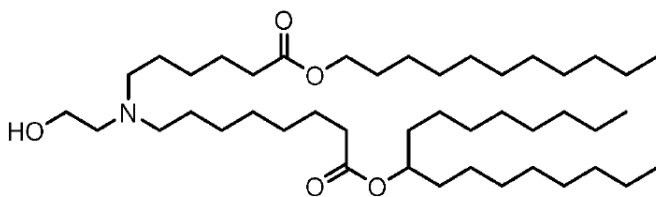
39. El método de cualquiera de los párrafos de método precedentes, en donde el sujeto
tiene 5 años de edad o menos, o 65 años de edad o más.

40. El método de cualquiera de los párrafos de método precedentes, que comprende
administrar al sujeto al menos dos dosis de la composición.

30 41. El método de cualquiera de los párrafos de método precedentes, en donde se
producen niveles detectables del antígeno del coronavirus en el suero del sujeto 1-72 horas
luego de la administración del ARN o composición que comprende el ARN.

42. El método de cualquiera de los párrafos de método precedentes, en donde se produce un título de anticuerpo neutralizante de al menos 100 NU/ml, al menos 500 NU/ml o al menos 1000 NU/ml en el suero del sujeto a 1-72 horas luego de la administración del ARN o composición que comprende el ARN.
- 5 43. Una composición inmunizante que comprende: una nanopartícula lipídica que comprende (a) un ARN mensajero que comprende un marco de lectura abierto (ORF) que tiene al menos 90 %, al menos 95 %, al menos 98 % o 100 % de identidad con la secuencia de la SEQ ID NO: 28 y (b) una mezcla de lípidos que comprende 0.5-15 % en moles de lípido modificado con PEG, 5-25 % en moles de lípido no catiónico, 25-55 % en moles de
10 esterol y 20-60 % en moles de lípido catiónico ionizable.
44. Una composición inmunizante que comprende: una nanopartícula lipídica que comprende (a) un ARN mensajero que comprende una secuencia que tiene al menos 90 %, al menos 95 %, al menos 98 % o 100 % de identidad con la secuencia de SEQ ID NO: 27 y (b) una mezcla de lípidos que comprende 0.5-15 % en moles de lípidos modificados con
15 PEG, 5-25 % en moles de lípidos no catiónicos, 25-55 % en moles de esterol y 20-60 % en moles de lípidos catiónicos ionizables.
45. Una composición inmunizante que comprende:
(a) un primer ácido ribonucleico (ARN) que comprende un marco de lectura abierto (ORF) que codifica un antígeno de coronavirus capaz de inducir una respuesta inmune, tal como
20 una respuesta de anticuerpos neutralizantes, a un SARS-CoV-2; y
(b) un segundo ácido ribonucleico (ARN) que comprende un marco de lectura abierto (ORF) que codifica un antígeno de coronavirus capaz de inducir una respuesta inmune, como una respuesta de anticuerpos neutralizantes, a un SARS-CoV-2, en donde el ORF del primer ARN es diferente del ORF del segundo ARN.
- 25 46. La composición inmunizante del párrafo 45 que comprende además una nanopartícula lipídica que comprende una mezcla de lípidos.
47. La composición inmunizante del párrafo 46, en donde la mezcla de lípidos comprende un lípido modificado con PEG, un lípido no catiónico, un esterol, un lípido catiónico ionizable o cualquier combinación de estos.
- 30 48. La composición inmunizante del párrafo 47, en donde la mezcla de lípidos comprende 0.5-15 % en moles de lípido modificado con PEG; 5-25 % en moles de lípidos no catiónicos; 25-55 % en moles de esterol; y 20-60 % en moles de lípido catiónico ionizable.

49. La composición inmunizante del párrafo 46 o 47, en donde el lípido modificado con PEG es 1,2 dimiristoil-sn-glicerol, metoxipolietilenglicol (PEG2000 DMG), el lípido no catiónico es 1,2 diestearoil-sn-glicero-3- fosfocolina (DSPC), el esteroles es colesterol y el lípido catiónico ionizable tiene la estructura del Compuesto 1:



(Compuesto 1).

50. La composición inmunizante de cualquiera de los párrafos 45-49, en donde el antígeno de coronavirus codificado por el ORF del primer ARN comprende una secuencia de aminoácidos que tiene al menos 80 %, al menos 85 %, al menos 90 % o al menos 95 % de identidad con la secuencia de aminoácidos de cualquiera de las SEQ ID NO: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83 y 85.

51. La composición inmunizante de cualquiera de los párrafos 45-49, en donde el ORF del primer ARN comprende una secuencia de nucleótidos que tiene al menos 80 %, al menos 85 %, al menos 90 % o al menos 95 % de identidad con la secuencia de aminoácidos de cualquiera de las SEQ ID NO: 3, 7, 10, 13, 16, 19, 22, 25, 28, 31, 48, 50, 52, 54, 56, 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82 y 84.

52. El ARN de cualquiera de los párrafos 1-13, en donde el ORF codifica un antígeno del SARS-CoV-2.

53. El ARN del párrafo 52, en donde el antígeno del SARS-CoV-2 es una proteína estructural.

54. El ARN del párrafo 53, en donde la proteína estructural se selecciona del grupo que consiste en: una proteína de pico (S), una proteína de membrana (M), una proteína de envoltura (E) y una proteína de nucleocápside (NC).

55. El ARN del párrafo 54, en donde la proteína estructural es una proteína S, opcionalmente una forma de prefusión estabilizada de una proteína S.

56. El ARN del párrafo 55, en donde la proteína S es una variante de la proteína S, en relación con una proteína S que comprende la secuencia de aminoácidos de SEQ ID NO: 32.

57. El ARN del párrafo 56, en donde la variante de la proteína S comprende una reversión

de un sitio de escisión polibásico a un único sitio de escisión básico.

58. El ARN del párrafo 56, en donde la variante de la proteína S comprende una eliminación de una secuencia señal polibásica ER/Golgi (KXHXX-COOH) en la cola carboxi de la variante de la proteína S.

5 59. El ARN de los párrafos 57-58, en donde la proteína S comprende una mutación estabilizadora de prolina doble.

60. El ARN de los párrafos 57-58, en donde la proteína S comprende un sitio de escisión de proteasa modificado para estabilizar la proteína.

10 61. El ARN de los párrafos 57-60, en donde la proteína S comprende una eliminación de la cola citoplasmática.

62. El ARN de los párrafos 57-61, en donde la proteína S comprende un armazón de pliegue.

15 63. El ARN del párrafo 57, en donde la proteína S comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de cualquiera de las SEQ ID NO: 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 33, 34, 35, 47, 49, 59, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83 u 85.

64. El ARN del párrafo 58, en donde la proteína estructural es una proteína M.

65. El ARN del párrafo 64, en donde la proteína M comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 81.

20 66. El ARN del párrafo 65, en donde la proteína M comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 81.

67. El ARN de cualquiera de los párrafos 57-66, en donde el ORF comprende la secuencia de SEQ ID NO: 80.

25 68. El ARN de cualquiera de los párrafos 57-67, en donde el ARN comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 95.

69. El ARN del párrafo 68, en donde el ARN comprende la secuencia de SEQ ID NO: 95.

70. El ARN del párrafo 54, en donde la proteína estructural es una proteína E.

30 71. El ARN del párrafo 70, en donde la proteína E comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 83.

72. El ARN del párrafo 71, en donde la proteína E comprende una secuencia que tiene

al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 83.

73. El ARN de cualquiera de los párrafos 70-72, en donde el ORF comprende la secuencia de SEQ ID NO: 82.

5 74. El ARN de cualquiera de los párrafos 70-73, en donde el ARN comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 96.

75. El ARN del párrafo 74, en donde el ARN comprende la secuencia de SEQ ID NO: 96.

76. El ARN del párrafo 54, en donde la proteína estructural es una proteína NC.

10 77. El ARN del párrafo 76, en donde la proteína NC comprende una secuencia que tiene al menos un 80 % de identidad con la secuencia de SEQ ID NO: 85.

78. El ARN del párrafo 77, en donde la proteína NC comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 85.

15 79. El ARN de cualquiera de los párrafos 76-78, en donde el ORF comprende la secuencia de SEQ ID NO: 84.

80. El ARN de cualquiera de los párrafos 76-78, en donde el ARN comprende una secuencia que tiene al menos el 85 %, al menos el 90 %, al menos el 95 % o al menos el 98 % de identidad con la secuencia de SEQ ID NO: 97.

20 81. El ARN del párrafo 80, en donde el ARN comprende la secuencia de SEQ ID NO: 97.

82. El ARN del párrafo 53, en donde el antígeno del SARS-CoV-2 es una proteína de fusión.

83. El ARN del párrafo 82, en donde la proteína de fusión comprende un polipéptido de SARS-CoV-2 y un polipéptido de un virus diferente.

25 84. Un ácido ribonucleico mensajero (ARNm) que comprende un marco de lectura abierto (ORF) que comprende una secuencia de nucleótidos que tiene al menos 80 % de identidad con la secuencia de nucleótidos la SEQ ID NO: 106.

85. Un ácido ribonucleico mensajero (ARNm) que comprende una secuencia de nucleótidos que tiene al menos 80 % de identidad con la secuencia de nucleótidos la SEQ
30 ID NO: 105.

86. Un ácido ribonucleico mensajero (ARNm) que comprende un marco de lectura abierto (ORF) que comprende una secuencia de nucleótidos que tiene al menos 95 % de identidad

con la secuencia de nucleótidos la SEQ ID NO: 106.

87. Un ácido ribonucleico mensajero (ARNm) que comprende una secuencia de nucleótidos que tiene al menos 95 % de identidad con la secuencia de nucleótidos la SEQ ID NO: 105.

5 88. Un ácido ribonucleico mensajero (ARNm) que comprende un marco de lectura abierto (ORF) que comprende una secuencia de nucleótidos que tiene al menos 99 % de identidad con la secuencia de nucleótidos la SEQ ID NO: 106.

89. Un ácido ribonucleico mensajero (ARNm) que comprende una secuencia de nucleótidos que tiene al menos 99 % de identidad con la secuencia de nucleótidos la SEQ
10 ID NO: 105.

LISTADO DE SECUENCIAS

Debe entenderse que cualquiera de las secuencias de ARNm descritas en el presente documento puede incluir una 5' UTR y/o una 3' UTR. Las secuencias de UTR se pueden
15 seleccionar de las siguientes secuencias, o se pueden utilizar otras secuencias de UTR conocidas. También debe entenderse que cualquiera de las construcciones de ARNm descritas en el presente documento puede comprender además una cola y/o cubierta de poli(A) (p. ej., 7mG(5')ppp(5')NlmpNp). Además, aunque muchos de los ARNm y las secuencias de antígeno codificadas descritas en este documento incluyen un péptido señal
20 y/o una etiqueta peptídica (p. ej., etiqueta His C-terminal), debe entenderse que la etiqueta peptídica y/o péptida señal indicada puede sustituirse por un péptido señal y/o etiqueta de péptido diferente, o se puede omitir el péptido señal y/o etiqueta de péptido.

5' UTR: GGGAAAUAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGAGCCACC (SEQ
25 ID NO: 36)

5' UTR:
GGGAAAUAGAGAGAAAAGAAGAGUAAGAAGAAAUUAAGACCCCGGCGCCGCCAC
C (SEQ ID NO: 2)

3' UTR:
30 UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCC
AGCC
CCUCCUCCCCUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGG

GCGGC (SEQ ID NO: 37)

3'

UTR:

UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCUUCUUGCCCCUUGGGCCUCCCCC
AGCC

5 CCUCCUCCCCUUCUGACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGG
GCGGC (SEQ ID NO: 4)

Tabla 1

Proteína de pico (S) del SARS-CoV-2		
La SEQ ID NO: 30 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 31 y 3' UTR SEQ ID NO: 4.		30
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUUAUAGACCCCGGCGCCGCCACC	2.
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU	31

	ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCUUGCAACGGCGUGGAGGGCUUCAACUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC	
--	--	--

CACCAACGGCGUGGGGCUACCAGCCCUACCGGG
 UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC
 CCCAGCCACCGUGUGUGGGCCCAAGAAGAGCA
 CCAACCUGGUGAAGAACAAGUGCGUGAACUUC
 AACUUCAACGGCCUUACCGGCACCGGCGUGCU
 GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC
 AGCAGUUCGGCCGGGACAUCGCCGACACCACC
 GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU
 CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG
 UGAGCGUGAUCACCCAGGCACCAACACCAGC
 AACCAGGUGGCCGUGCUGUACCAGGACGUGAA
 CUGCACCGAGGUGCCCGUGGCCAUCCACGCC
 GACCAGCUGACACCCACCUGGCGGGUCUACAG
 CACCGGCAGCAACGUGUUCAGACCCGGGCGG
 GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC
 AGCUACGAGUGCGACAUCCCCAUCCGGCGCCGG
 CAUCUGUGCCAGCUACCAGACCCAGACCAAUU
 CACCCCGGAGGGCAAGGAGCGUGGCCAGCCA
 GAGCAUCAUCGCCUACACCAUGAGCCUGGGCG
 CCGAGAACAGCGUGGCCUACAGCAACAACAGC
 AUCGCCAUCCCCACCAACUUCACCAUCAGCGU
 GACCACCGAGAUUCUGCCCGUGAGCAUGACCA
 AGACCAGCGUGGACUGCACCAUGUACAUCUGC
 GGCGACAGCACCGAGUGCAGCAACCUGCUGCU
 GCAGUACGGCAGCUUCUGCACCCAGCUGAACC
 GGGCCCUGACCGGCAUCGCCGUGGAGCAGGA
 CAAGAACACCCAGGAGGUGUUCGCCAGGUGA
 AGCAGAUCAACAAGACCCCUCCAUCAAGGAC
 UUCGGCGGCUUCAACUUCAGCCAGAUCUGCC
 CGACCCAGCAAGCCCAGCAAGCGGAGCUUCA
 UCGAGGACCUGCUGUUCAACAAGGUGACCCUA
 GCCGACGCCGGCUUCAUCAAGCAGUACGGCGA

	CUGCCUCGGCGACAUAGCCGCCCGGGACCUGA UCUGCGCCCAGAAGUUCAACGGCCUGACCGUG CUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCA UCACCAGCGGCUGGACUUUCGGCGCUGGAGC CGCUCUGCAGAUCCCCUUCGCCAUGCAGAUGG CCUACCGGUUCAACGGCAUCGGCGUGACCCAG AACGUGCUGUACGAGAACCAGAAGCUGAUCGC CAACCAGUUCAACAGCGCCAUCGGCAAGAUC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUG GGCAAGCUGCAGGACGUGGUGAACCAGAACGC CCAGGCCCUGAACACCCUGGUGAAGCAGCUGA GCAGCAACUUCGGCGCCAUCAGCAGCGUGCUG AACGACAUCCUGAGCCGGCUGGACAAGGUGGA GGCCGAGGUGCAGAUCGACCGGCUGAUCACU GGCCGGCUGCAGAGCCUGCAGACCUACGUGAC CCAGCAGCUGAUCCGGGCCCGCCGAGAUUCGG GCCAGCGCCAACCUGGCCGCCACCAAGAUGAG CGAGUGCGUGCUGGGCCAGAGCAAGCGGGUG GACUUCUGC GGCAAGGGCUACCACCUGAUGAG CUUUC C C C CAGAGCGCACCCACGGAGUGGUGU UCCUGCAGUGACCUACGUGCCCGCCCAGGAG AAGAACUUCACCACCGCCCCAGCCAUCUGCCA CGACGGCAAGGCCACUUC C C C C G G GAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGU GACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAAC UGCGACGUGGUGAUCGGCAUCGUGAACAACAC CGUGUACGAUCCCCUGCAGCCCGAGCUGGACA GCUUCAAGGAGGAGCUGGACAAGUACUUCAAG AAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACA	
--	--	--

	UCCAGAAGGAGAUUCGAUCGGCUGAACGAGGUG GCCAAGAACCUGAACGAGAGCCUGAUCGACCU GCAGGAGCUGGGCAAGUACGAGCAGUACAUCA AGUGGCCCUGGUACAUCUGGCUGGGCUUCAU CGCCGGCCUGAUCGCCAUCGUGAUGGUGACCA UCAUGCUGUGCUGCAUGACCAGCUGCUGCAGC UGCCUGAAGGGCUGUUGCAGCUGCGGCAGCU GCUGCAAGUUCGACGAGGACGACAGCGAGCCC GUGCUGAAGGGCGUGAAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLVNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKS NLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPI	32

	GAGICASYQTQTNSPRRRARSVASQSIIAYTMSLGA ENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTEC SNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFN SAIGKIQDSLSTASALGKLQDV VNQNAQALNTLVKQLSSNFGAISSVLNDILSRDKV EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPRE GVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCC MTSCCCLKGCCSCGSCCKFDEDDSEPV LKGVKL HYT	
Cola de PoliA	100 nt	
Variante 1 de la proteína S del SARS-CoV-2		
La SEQ ID NO: 1 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 3 y 3' UTR SEQ ID NO: 4.		1
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el	AUGUUCGUCUUCUGUCUUGCUGCCGCUGG UGUCGAGCCAGUGCGUGAACCUCACCACAAGG ACGCAGCUCCACCGGCCUACACGAACAGCUU CACGCGCGGCGUGUACUACCCCGACAAGGUGU	3

codón de detención)	UCCGGUCGUCGGUCCUCCACUCCACGCAGGAC CUCUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUCUCCGGGACGAACG GGACGAAGCGGUUCGACAACCCGGUCCUCCCG UUCAACGACGGCGUCUACUUCGCGAGCACGGA GAAGUCGAACAUCAUCCGGGGCUGGAUCUUCG GCACGACCCUGGACUCGAAGACCCAGUCCCUA CUUAUCGUGAACAACGCCACCAACGUCGUCAU CAAGGUCUGCGAGUUCAGUUCUGCAACGACC CCUUCCUCGGCGUCUACUACCACAAGAACAAC AAGUCGUGGAUGGAGUCGGAGUUCGGGUGU ACAGCUCGGCGAACAACUGCACCUUCGAGUAC GUGUCGCAGCCGUUCCUCAUGGACCUCGAGG GCAAGCAGGGUAACUUCAAGAACCUGCGCGAG UUCGUCUUCAAGAACAUCGACGGCUACUUCAA GAUCUACUCCAAGCACACGCCCAUCAACCUGG UCCGCGACCUCGCGCAAGGCUUCUCCGCCUC GAGCCUCUGGUCGACCUGCCGAUCGGCAUCAA CAUCACGAGGUUCCAGACGCUCCUGGCGCUGC ACCGGUCGUACCUGACGCCAGGCGACUCCUCC UCGGGCUGGACAGCAGGCGCGGCUGCCUACU ACGUCGGGUACCUGCAGCCCCGCACGUUCCUC CUGAAGUACAACGAGAACGGCACUAUCACGGA CGCCGUCGACUGCGCCCUGGACCCACUGUCG GAGACGAAGUGCACGCUGAAGUCGUUACCGU GGAGAAGGGUAUCUACCAGACCUCCAACUUC GGGUCCAGCCGACGGAGUCGAUCGUGCGGUU CCCCAACAUACGAACCUGUGCCCCUUCGGUG AGGUCUUAACGCCACCCGGUUCGCGUCGGUC UACGCGUGGAACCGUAAGCGCAUCUCGAACUG CGUGGCGGACUACUCCGUCCUCUACAACAGCG CGUCCUUCAGCACCUUCAAGUGCUACGGCGUC	
----------------------------	---	--

	AGCCCCACGAAGCUGAACGACCUCUGCUUCAC CAACGUCUACGCAGACUCCUUCGUGAUCCGGG GUGACGAGGUGCGACAGAUCGCCCCUGGUCA GACCGGGAAGAUCGCCGACUACAACUACAAGC UGCCCGACGACUUCACCGGCUGCGUGAUCGC GUGGAACAGCAACAACCUGGACUCCAAGGUCG GAGGUAACUACAACUACCUCUACCGGCUGUUC CGCAAGUCCAACCUGAAGCCGUUCGAGCGGGA CAUCUCCACGGAGAUCUACCAAGCCGGCUCGA CCCUUGUAACGGGGUGGAGGGGUUCAACUG CUACUUCCCACUGCAGUCCUACGGGUUCCAGC CCACCAACGGGGUCGGGUACCAGCCGUACCGC GUGGUGGUCCUGUCCUUCGAGCUGCUGCACG CGCCAGCCACGGUGUGCGGGCCAAAGAAGAGC ACGAACCUGGUCAAGAACAAGUGCGUCAACUU CAACUUCAACGGCCUGACGGGGACAGGGGUCC UCACGGAGUCGAACAAGAAGUUCCUGCCGUUC CAGCAGUUCGGCCGUGACAUCGCAGACACGAC UGACGCCGUCCGCGACCCUCAGACCCUCGAGA UCCUCGACAUCACCCCGUGCUCGUUCGGCGGA GUGAGCGUCAUCACCCCGGGGACCAACACAUC GAACCAGGUGGCGGUCCUGUACCAGGACGUCA ACUGCACGGAGGUCCCUGUGGCGAUCCACGCC GACCAGCUCACGCCACCUGGCGCGUCUACUC CACCGGGUCCAACGUGUUCCAGACCCGCGCAG GCUGCCUGAUCGGGGCCGAGCACGUCAACAAC AGCUACGAGUGCGACAUCCCAUCGGAGCGGG CAUCUGCGCCAGCUACCAGACGCAGACGAACU CUCCAAGGCGCGCUCGUAGCGUGGCCUCCCA GUCCAUCAUCGCGUACACGAUGUCCCUUGGGG CCGAGAACUCGGUCGCAUACAGCAACAACUCC AUCGCCAUCCCCACCAACUUCACGAUCUCGGU	
--	--	--

	CACCACCGAGAUCCUCCCGGUCAGCAUGACGA AGACGUCGGUGGACUGCACCAUGUACAUCUGC GGGGACAGCACGGAGUGCUCGAACCUGCUCCU GCAGUACGGGAGCUUCUGCACCCAGCUGAACA GGGCGCUGACGGGGAUCGCGGUGGAGCAGGA CAAGAACACCCAGGAGGUGUUCGCGCAGGUGA AGCAGAUUCUACAAGACGCCUCCAAUCAAGGAC UUCGGCGGGUUAACUUCUCGCAGAUCCUCCC CGACCCGUCCAAGCCGUCGAAGCGGUCGUUCA UCGAGGACCUGCUCUUAACAAGGUGACGUUG GCCGACGCGGGCUUCAUCAAGCAGUACGGGGA CUGCCUUGGGGACAUCGCUGCCCGCGACCUC UCUGCGCCCAGAAGUUAACGGGCUGACUGUG CUCCCGCCCCUGCUGACGGACGAGAUGAUCGC GCAGUACACGUCCGCGCUGCUCGCUGGAACGA UCACCUCCGGGUGGACCUUCGGCGCUGGAGC GGCUCUGCAGAUCCCGUUCGCGAUGCAGAUG GCGUACCGGUUAACGGCAUCGGGGUGACCCA GAACGUCCUCUACGAGAACCAGAAGCUGAUCG CCAACCAGUUAACUCCGCGAUCGGCAAGAUC CAGGACUCGCUGAGCUCCACGGCUUCCGCCCU CGGGAAGCUUCAGGACGUGGUGAACCAAGACG CCCAGGCCCUAACACCCUGGUGAAGCAGCUG AGCUCGAACUUCGGCGCCAUCUCGAGCGUGCU CAACGACAUCCUGAGCCGUCUGGACCCUCCG AGGCGGAGGUGCAGAUCGACCGGCUCAUCACG GGCCGGCUUCAGUCCCUGCAGACGUACGUGAC CCAGCAGCUCAUACGGGCGGCGGAGAUACGCG CCUCCGCCAACCUGGCCGCGACGAAGAUGUCC GAGUGCGUCCUCGGACAGAGCAAGCGCGUGG ACUUCUGCGGCAAGGGGUACCACCUCAUGAGC UUUCCCCAGUCGGCUCCUCACGGGGUCGUCU	
--	---	--

	UCCUGCACGUGACGUACGUCCCCGGCGCAGGA GAAGAACUUCACCACCGCCCCAGCGAUCUGCC ACGACGGGAAGGCGCACUUCCCGCGCGAGGG CGUCUUCGUCUCCAACGGGACCCACUGGUUCG UCACCCAGCGGAACUUCUACGAGCCGCAGAUC AUCACGACCGACAACACGUUCGUAUCCGGGAA CUGCGACGUCGUCAUCGGCAUCGUCAACAACA CGGUCUACGACCCACUGCAGCCGGAGCUGGAC UCGUUCAAGGAGGAGCUGGACAAGUAUUUCAA GAACCACACCUCGCCCCGACGUGGACCUGGGCG ACAUCAGCGGGAUCAACGCGUCGGUCGUGAAC AUCCAGAAGGAGAUCGACCGACUGAACGAGGU CGCCAAGAACCUGAACGAGUCCCUGAUCGACC UGCAAGAGCUCGGCAAGUACGAGCAGUACAUC AAGUGGCCUUGGUACAUCUGGCUCGGCUUCAU CGCGGGGCUGAUCGCCAUCGUGAUGGUCACC AUCAUGUUGUGCUGCAUGACCUCUGCUGCUC GUGCCUCAAGGGGUGCUGCAGCUGCGGGUCC UGCUGCAAGUUCGACGAGGACGACUCGGAGCC GGUCCUCAAGGGCGUCAAGCUCCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNE	5

	NGTITDAVDCALDPLSETKCTLKSTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSP TKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKS NLKP FERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGA EHVNNSECDIPI GAGICASYQTQTNSPRRARSVASQSIIAYTMSLGA ENSVAYSNN SIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTEC SNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKR SFIEDLLFNKVT LADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPPLL TDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFN SAIGKIQDSL SSTA SALGKLQDV VNQNAQALNTLVKQLSSNFGA ISSVLNDILSRLDPP EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICH DGKAHFPRE GV FVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIA GLIAIVMVTIMLCC MTSCCCLKGCCSCGSCCKFDEDDSEPV LKGVKL HYT	
Cola de PoliA	100 nt	
Variante 2 de la proteína del SARS-CoV-2 S		

La SEQ ID NO: 6 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 7 y 3' UTR SEQ ID NO: 4.		6
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCAGG ACCCAGCUGCCGCCUGCCUACACCAACAGCUU CACCCGCGGUGUGUACUACCCCGACAAGGUGU UCAGGUCCAGCGUGCUGCACAGCACCCAGGAC CUGUUCCUCCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACACUCGACAGCAAGACCCAGAGCCUG CUGAUCGUGAACAACGCCACCAACGUGGUGAU CAAGGUGUGCGAAUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGCGUGUA CAGCAGCGCCAACAACUGCACCUUCGAGUACG UGAGCCAGCCCUUCCUGAUGGACCUGGAGGG CAAGCAGGGCAAUUUCAAGAACCUGAGGGAGU UCGUGUUCAAGAACAUCGACGGCUACUUCAAG AUCUACAGCAAGCACACGCCCAUCAACCUGGU GCGGGACUUGCCCCAGGGCUUCAGCGCCCUG GAGCCCUUAGUGGACCUGCCUAUCGGCAUCAA CAUCACCCGGUUCCAGACCCUGCUGGCCCUGC ACCGGAGCUACCUGACUCCCGGCGACAGCAGC	7

	UCCGGGUGGACUGCCGGUGCUGCCGCCUACU ACGUGGGGUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGAUCCACUGAGC GAGACCAAGUGCACCCUGAAGAGCUUCACCGU GGAGAAGGGCAUCUACCAGACCAGCAACUUC GGGUGCAGCCCACCGAGAGCAUCGUGAGGUU CCCCACAUCACCAACCUGUGCCCUUUCGGCG AGGUGUUCAACGCCACCCGCUUCGCCUCCGUG UACGCCUGGAACAGGAAGAGGAUCAGCAACUG CGUGGCCGACUACAGCGUGCUGUACAACAGCG CCAGCUUCUCCACCUUCAAGUGCUACGGCGUG AGCCCAACCAAGCUGAACGACCUGUGCUUUAC CAACGUGUACGCCGAUAGCUUCGUGAUCCGCG GCGACGAAGUGCGGCAGAUCGCUCCUGGGCA GACCGGAAAGAUCGCCGACUACAACUACAAGC UGCCCCGACGACUUCACCGGGUGCGUGAUCGC UUGGAACAGCAACAACCUGGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGCGA CAUCUCCACCGAGAUCUACCAGGCCGGCUCCA CACCCUGCAACGGCGUGGAGGGCUUCAAUCGC UACUUUCCCCUGCAGUCCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCAUAACGCG UGGUGGUGCUGUCCUUCGAGCUGCUGCACGC UCCCGCCACCGUUUGCGGCCCAAGAAGUCCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGUCUCACGGGCACCGGGGUGCU GACCGAGAGCAACAAGAAGUUCUGCCCUUUC AGCAGUUCGGCAGGGACAUCGCCGACACCACA GACGCCGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCGUGCAGCUUCGGCGGC	
--	--	--

	GUGAGCGUGAUCACGCCCGGCACCAACACCAG CAACCAGGUGGCCGUGCUGUACCAGGACGUGA ACUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACUCCCACCUGGCGCGUGUAUAG CACCGGCAGCAACGUGUUCCAGACACGGGCCG GCUGCCUGAUCGGCGCCGAGCACGUGAACAAC UCCUACGAGUGCGACAUCCCCAUCCGGCGCUGG CAUCUGCGCCAGCUACCAGACCCAGACCAACA GCCCCAGACGGGCCAGGUCCGUGGCUUCCCA GAGCAUCAUCGCCUACACCAUGUCCCUGGGCG CCGAGAACAGCGUGGCCUACAGCAACAACUCC AUCGCCAUCCCCACCAACUUCACCAUCAGCGU GACCACCGAGA UCCUGCCCGUGAGCAUGACCA AGACCUCGGUGGACUGCACCAUGUACAUCUGC GGCGACAGCACCGAGUGCAGCAACCUGCUGCU GCAGUACGGCAGCUUCUGCACCCAGCUGAACA GGGCCUGACCGGCAUCGCCGUGGAGCAGGA CAAGAACACCCAGGAGGUGUUCGCCCAGGUGA AGCAGAUUCUACAAGACUCCACCUAUCAAGGAC UUCGGCGGGUUAACUUCAGCCAGA UCCUCCC CGACCCCUCCAAGCCCAGCAAGCGGAGCUUCA UCGAGGACCUGCUGUUAACAAGGUGACCCUG GCUGACGCCGGCUUUAUCAAGCAGUACGGCGA CUGCCUUGGCGACAUCGCCGCCAGGGACCUGA UCUGCGCCCAGAAGUUAACGGCCUGACCGUG CUGCCGCCACUGCUGACCGACGAGAUGAUCGC CCAGUACACCUCUGCCCUGCUGGCCGGUACCA UCACCUCGGCUGGACA UUGGUGCUGGGCGC UGCGCUGCAGA UCCCCUUCGCCAUGCAGAUGG CCUACCGCUUAACGGCAUCGGGGUGACCCAG AACGUGCUGUACGAGAACCAGAAGCUGAUCGC CAACCAGUUAACAGCGCCAUCGGCAAGA UCC	
--	--	--

	AGGACAGCCUGAGCAGCACCGCCAGCGCUCUG GGCAAGCUGCAGGACGUGGUGAACCAGAACGC CCAGGCCCUGAACACCCUGGUGAAGCAGCUGU CCAGCAACUUCGGCGCCAUCAGCUCCGUGCUG AACGACAUCCUGAGCCGGCUGGAUCCACCAGA GGCCGAGGUGCAGAUCGACCGUCUGAUCACCG GUCGGCUGCAGAGCCUGCAGACCUACGUGACC CAGCAGCUGAUCCGCGCCGCGGAAAUCCGCGC CUCCGCCAACCUGGCCGCCACCAAGAUGUCCG AGUGCGUGCUGGGCCAGAGCAAGCGGGUGGA CUUCUGCGGCAAGGGCUACCACCUGAUGAGCU UCCCACAGAGCGCUCCCCACGGGGUAGUGUUC CUGCACGUGACCUACGUGCCCGCCCAGGAGAA GAACUUCACCACUGCACCCGCCAUCUGCCACG ACGGCAAGGCCACUUCCCUCGGGAGGGCGU GUUCGUGAGCAACGGCACCCACUGGUUCGUGA CCCAGAGGAACUUCUACGAGCCCCAGAUCAUC ACCACCGACAACACCUUCGUGUCCGGCAACUG CGACGUGGUGAUCGGCAUAGUGAACAACACCG UGUACGACCCACUGCAGCCCGAGCUGGACAGC UUCAAGGAGGAGCUGGACAAGUACUUCAAGAA CCACACCAGCCCAGACGUGGACCUGGGCGACA UCUCCGGCAUCAACGCCUCCGUGGUGAACAUC CAGAAGGAGAUCGACCGGCUGAACGAGGUGGC CAAGAACCUGAACGAGAGCCUGAUCGACCUGC AGGAGCUGGGGAAGUACGAGCAGUACAUAAG UGGCCUUGGUACAUCUGGCUGGGCUUCAUCG CCGGCCUGAUCGCCAUCGUGAUGGUGACCAUC AUGCUGUGCUGCAUGACCAGCUGCUGCAGCUG CCUGAAGGGCUGUUGCAGCUGCGGCAGCUGC UGCAAGUUCGACGAGGACGACAGCGAGCCCGU GCUGAAGGGCGUGAAGCUGCACUACACC	
--	--	--

3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUCCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLVNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSECDIPI GAGICASYQTQTNSPRRARSVASQSIIAYTMSLGA ENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTECNNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPLLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFNSAIGKIQDSLSTASALGKLQDV VNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPP	8

	EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPRE GVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCC MTSCCCLKGCCSCGSCCKFDEDDSEPVLKGVKL HYT	
Cola de PoliA	100 nt	
Variante 3 de la proteína SARS-CoV-2 S		
SEQ ID NO: 9 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 10 y 3' UTR SEQ ID NO: 4.		9
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NmpNp	
5' UTR	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC	10

	CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUC CAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAU CGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCU CGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA	
--	--	--

	CCCCUUGCAACGGCGUGGAGGGCUUCAACUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCCAGACCCGGGCCG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCCAUCCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCAAUU CACCCGGCAGCGGCGGCAGCGUGGCCAGCCA GAGCAUCAUCGCCUACACCAUGAGCCUGGGCG CCGAGAACAGCGUGGCCUACAGCAACAACAGC AUCGCCAUCCCCACCAACUUCACCAUCAGCGU GACCACCGAGAUUCUGCCCGUGAGCAUGACCA AGACCAGCGUGGACUGCACCAUGUACAUCUGC GGCGACAGCACCGAGUGCAGCAACCUGCUGCU GCAGUACGGCAGCUUCUGCACCCAGCUGAACC GGGCCUGACCGGCAUCGCCGUGGAGCAGGA CAAGAACACCCAGGAGGUGUUCGCCCAGGUGA AGCAGAUCAACAAGACCCCUCCAUCAAGGAC UUCGGCGGCUUCAACUUCAGCCAGAUCUGCC CGACCCAGCAAGCCCAGCAAGCGGAGCUUCA	
--	--	--

	UCGAGGACCUGCUGUUCAACAAGGUGACCCUA GCCGACGCCGGCUUCAUCAAGCAGUACGGCGA CUGCCUCGGCGACAUAAGCCGCCCGGGACCUGA UCUGCGCCCAGAAGUUCAACGGCCUGACCGUG CUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCA UCACCAGCGGCUGGACUUUCGGCGCUGGAGC CGCUCUGCAGAUCCCCUUCGCCAUGCAGAUGG CCUACCGGUUCAACGGCAUCGGCGUGACCCAG AACGUGCUGUACGAGAACCAGAAGCUGAUCGC CAACCAGUUCAACAGCGCCAUCGGCAAGAUGC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUG GGCAAGCUGCAGGACGUGGUGAACCAGAACGC CCAGGCCCUGAACACCCUGGUGAAGCAGCUGA GCAGCAACUUCGGCGCCAUCAGCAGCGUGCUG AACGACAUCCUGAGCCGGCUGGACCCUCCCGA GGCCGAGGUGCAGAUUCGACCGGCUGAUCACU GGCCGGCUGCAGAGCCUGCAGACCUACGUGAC CCAGCAGCUGAUCCGGGCGCCGAGAUUCGG GCCAGCGCCAACCUGGCCGCCACCAAGAUGAG CGAGUGCGUGCUGGGCCAGAGCAAGCGGGUG GACUUCUGCGGCAAGGGCUACCACCUGAUGAG CUUUCCCCAGAGCGCACCCACGGAGUGGUGU UCCUGCACGUGACCUACGUGCCCGCCCAGGAG AAGAACUUCACCACCGCCCCAGCCAUCUGCCA CGACGGCAAGGCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGU GACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAAC UGCGACGUGGUGAUCGGCAUCGUGAACAACAC CGUGUACGAUCCCCUGCAGCCCGAGCUGGACA GCUUCAAGGAGGAGCUGGACAAGUACUUCAAG	
--	---	--

	AAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACA UCCAGAAGGAGAUCGAUCGGCUGAACGAGGUG GCCAAGAACCUGAACGAGAGCCUGAUCGACCU GCAGGAGCUGGGCAAGUACGAGCAGGGCAGC GGCUACAUCCCCGAGGCCCCUAGAGACGGCCA GGCCUACGUGCGGAAGGACGGCGAGUGGGUG CUGCUGAGCACCUUCCUG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKS NLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPI GAGICASYQTQTNSPGSGGSVASQSIIAYTMSLGA	11

	ENSVAYSNNNSIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTEC SNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFNSAIGKIQDSLSTASALGKLQDV VNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPP EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICHDKAHFPRE GVFSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQSGSYIPEAPRDGQAYVRKDGGEWVLLS TFL	
Cola de PoliA	100 nt	
Variante 4 de la proteína SARS-CoV-2 S		
La SEQ ID NO: 12 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 13 y 3' UTR SEQ ID NO: 4.		12
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG	13

	GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU	
--	---	--

	GGCGACGAGGUGCGGCAGAU CGCACCCGGCC AGACAGGCAAGAU CGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAU CGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAU CUACCAAGCCGGCUCCA CCCUUGCAACGGCGUGGAGGGCUUCAA CUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG UGAGCGUGAUACCCCAGGCACCAACACCAGC AACCAGGUGGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCCG GUUGCCUGAU CGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCCAU CGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCAAUU CACCCCGGAGGGCAAGGAGCGUGGCCAGCCA GAGCAUCAUCGCCUACACCAUGAGCCUGGGCG CCGAGAACAGCGUGGCCUACAGCAACAACAGC AUCGCCAUCCCCACCAACUUCACCAUCAGCGU GACCACCGAGAUUCUGCCCGUGAGCAUGACCA AGACCAGCGUGGACUGCACCAUGUACAUCUGC	
--	---	--

GGCGACAGCACCGAGUGCAGCAACCUGCUGCU
 GCAGUACGGCAGCUUCUGCACCCAGCUGAACC
 GGGCCCUGACCGGCAUCGCCGUGGAGCAGGA
 CAAGAACACCCAGGAGGUGUUCGCCCAGGUGA
 AGCAGAUUCUACAAGACCCCUCCCAUCAAGGAC
 UUCGGCGGCUUCAACUUCAGCCAGAUCUGCC
 CGACCCAGCAAGCCCAGCAAGCGGAGCUUCA
 UCGAGGACCUGCUGUUCAACAAGGUGACCCUA
 GCCGACGCCGGCUUCAUCAAGCAGUACGGCGA
 CUGCCUCGGCGACAUAAGCCGCCCGGGACCUGA
 UCUGCGCCCAGAAGUUCAACGGCCUGACCGUG
 CUGCCUCCCCUGCUGACCGACGAGAUGAUCGC
 CCAGUACACCAGCGCCCUGUUAGCCGGAACCA
 UCACCAGCGGCUGGACUUUCGGCGCUGGAGC
 CGCUCUGCAGAUCCCCUUCGCCAUGCAGAUGG
 CCUACCGGUUCAACGGCAUCGGCGUGACCCAG
 AACGUGCUGUACGAGAACCAGAAGCUGAUCGC
 CAACCAGUUCAACAGCGCCAUCGGCAAGAUC
 AGGACAGCCUGAGCAGCACCGCUAGCGCCCUG
 GGCAAGCUGCAGGACGUGGUGAACCAGAACGC
 CCAGGCCUGAACACCCUGGUGAAGCAGCUGA
 GCAGCAACUUCGGCGCCAUCAGCAGCGUGCUG
 AACGACAUCUGAGCCGGCUGGACCCUCCCGA
 GGCCGAGGUGCAGAUCGACCGGCUGAUCACU
 GGCCGGCUGCAGAGCCUGCAGACCUACGUGAC
 CCAGCAGCUGAUCCGGGCCCGCGAGAUCGG
 GCCAGCGCCAACCUUGGCCGCCACCAAGAUGAG
 CGAGUGCGUGCUGGGCCAGAGCAAGCGGGUG
 GACUUCUGCGGCAAGGGCUACCACCUGAUGAG
 CUUUCCCCAGAGCGCACCCACGGAGUGGUGU
 UCCUGCACGUGACCUACGUGCCCGCCCAGGAG
 AAGAACUUCACCACCGCCCCAGCCAUCUGCCA

	CGACGGCAAGGCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGU GACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAAC UGCGACGUGGUGAUCGGCAUCGUGAACAACAC CGUGUACGAUCCCCUGCAGCCCGAGCUGGACA GCUUCAAGGAGGAGCUGGACAAGUACUUCAAG AAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACA UCCAGAAGGAGAUCAUCGGCUGAACGAGGUG GCCAAGAACCUGAACGAGAGCCUGAUCGACCU GCAGGAGCUGGGCAAGUACGAGCAGGGCAGC GGCUACAUCCCCGAGGCCCUAGAGACGGCCA GGCCUACGUGCGGAAGGACGGCGAGUGGGUG CUGCUGAGCACCUUCCUG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLVNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLF	14

	RKS NLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPI GAGICASYQTQTNSPRRARSVASQSIIAYTMSLGA ENSVAYSNNNSIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTEC SNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFNSAIGKIQDSLSTASALGKLQDV VNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPP EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPRE GVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQGGSGYIPEAPRDGQAYVRKDG EWVLLS TFL	
Cola de PoliA	100 nt	
Variante 5 de la proteína SARS-CoV-2 S		
La SEQ ID NO: 15 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 16 y 3' UTR SEQ ID NO: 4.		15
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAU AAGAGAGAAAAGAAGAGUAAGAAGAA	2

	AUAUAAGACCCCGGCGCCGCCACC	
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU	16

	UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAU CGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCCUGCAACGGCGUGGAGGGCUUCAAACUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGGCCG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCAUCGGCGCCGG	
--	--	--

	CAUCUGUGCCAGCUACCAGACCCAGACCAAUU CACCCGGCAGCGGCGGCAGCGUGGCCAGCCA GAGCAUCAUCGCCUACACCAUGAGCCUGGGCG CCGAGAACAGCGUGGCCUACAGCAACAACAGC AUCGCCAUCCCCACCAACUUCACCAUCAGCGU GACCACCGAGAUUCUGCCCGUGAGCAUGACCA AGACCAGCGUGGACUGCACCAUGUACAUCUGC GGCGACAGCACCGAGUGCAGCAACCUGCUGCU GCAGUACGGCAGCUUCUGCACCCAGCUGAACC GGGCCUGACCGGCAUCGCCGUGGAGCAGGA CAAGAACACCCAGGAGGUGUUCGCCCAGGUGA AGCAGAUUCUACAAGACCCCUCCCAUCAAGGAC UUCGGCGGCUUCAACUUCAGCCAGAUCUGCC CGACCCAGCAAGCCCAGCAAGCGGAGCUUCA UCGAGGACCUGCUGUUCAACAAGGUGACCCUA GCCGACGCCGGCUUCAUCAAGCAGUACGGCGA CUGCCUCGGCGACAUAGCCGCCCGGGACCUGA UCUGCGCCCAGAAGUUCAACGGCCUGACCGUG CUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAAGCCGGAACCA UCACCAGCGGCUGGACUUUCGGCGCUGGAGC CGCUCUGCAGAUCCCCUUCGCCAUGCAGAUGG CCUACCGGUUCAACGGCAUCGGCGUGACCCAG AACGUGCUGUACGAGAACCAGAAGCUGAUCGC CAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUG GGCAAGCUGCAGGACGUGGUGAACCAGAACGC CCAGGCCCUGAACACCCUGGUGAAGCAGCUGA GCAGCAACUUCGGCGCCAUCAGCAGCGUGCUG AACGACAUCUGAGCCGGCUGGACCCUCCCGA GGCCGAGGUGCAGAUUCGACCGGCUGAUCACU GGCCGGCUGCAGAGCCUGCAGACCUACGUGAC	
--	---	--

	CCAGCAGCUGAUCCGGGCCCGAGAUUCGG GCCAGCGCCAACCUGGCCGCCACCAAGAUGAG CGAGUGCGUGCUGGGCCAGAGCAAGCGGGUG GACUUCUGCGGCAAGGGCUACCACCUGAUGAG CUUUCCCCAGAGCGCACCCCCACGGAGUGGUGU UCCUGCACGUGACCUACGUGCCCGCCCAGGAG AAGAACUUCACCACCGCCCCAGCCAUCUGCCA CGACGGCAAGGCCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGU GACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAAC UGCGACGUGGUGAUCGGCAUCGUGAACAACAC CGUGUACGAUCCCCUGCAGCCCGAGCUGGACA GCUUCAAGGAGGAGCUGGACAAGUACUUCAAG AAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACA UCCAGAAGGAGAUCGAUCGGCUGAACGAGGUG GCCAAGAACCUGAACGAGAGCCUGAUCGACCU GCAGGAGCUGGGCAAGUACGAGCAGUACAUCA AGUGGCCCUGGUACAUCUGGCUGGGCUUCAU CGCCGGCCUGAUCGCCAUCGUGAUGGUGACCA UCAUGCUGUGCUGCAUGACCAGCUGCUGCAGC UGCCUGAAGGGCUGUUGCAGCUGCGGCAGCU GCUGCAAGUUCGACGAGGACGACAGCGAGCCC GUGCUGAAGGGCGUGAAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI	17

	FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKS NLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSECDIPI GAGICASYQTQTNSPGSGGSVASQSIIAYTMSLGA ENSVAYSNNNSIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTEC SNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFNSAIGKIQDSLSTASALGKLQDV VNQNAQALNTLVKQLSSNFGAISSVLNDILSRDPP EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPRE GVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCC	
--	---	--

	MTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKL HYT	
Cola de PoliA	100 nt	
Variante 6 de la proteína SARS-CoV-2 S		
La SEQ ID NO: 18 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 19 y 3' UTR SEQ ID NO: 4.		18
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAAUAGAGAGAAAAGAAGAGUAAGAAGAA AUUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUCCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAUCAACCUGG	19

	UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUCACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAU CGCACCCGGCC AGACAGGCAAGAU CGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCCUGCAACGGCGUGGAGGGCUUCAAUCUG UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGCCUUAACCGGCACCGGCGUGCU	
--	--	--

	GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGCG UGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCCG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCCAUCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCAAUU CACCCCGGAGGGCAAGGAGCGUGGCCAGCCA GAGCAUCAUCGCCUACACCAUGAGCCUGGGCG CCGAGAACAGCGUGGCCUACAGCAACAACAGC AUCGCCAUCCCCACCAACUUCACCAUCAGCGU GACCACCGAGAUUCUGCCCGUGAGCAUGACCA AGACCAGCGUGGACUGCACCAUGUACAUCUGC GGCGACAGCACCGAGUGCAGCAACCUGCUGCU GCAGUACGGCAGCUUCUGCACCCAGCUGAACC GGGCCUGACCGGCAUCGCCGUGGAGCAGGA CAAGAACACCCAGGAGGUGUUCGCCCAGGUGA AGCAGAUCUACAAGACCCCUCCCAUCAAGGAC UUCGGCGGCUUCAACUUCAGCCAGAUCUGCC CGACCCAGCAAGCCCAGCAAGCGGAGCUUCA UCGAGGACCUGCUGUUCAACAAGGUGACCCUA GCCGACGCCGGCUUCAUCAAGCAGUACGGCGA CUGCCUCGGCGACAUAGCCGCCCGGGACCUGA UCUGCGCCCAGAAGUUCAACGGCCUGACCGUG CUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUUGUAGCCGGAACCA UCACCAGCGGCUGGACUUUCGGCGCUGGAGC	
--	--	--

	CGCUCUGCAGAUCCCCUUCGCCAUGCAGAUGG CCUACCGGUUCAACGGCAUCGGCGUGACCCAG AACGUGCUGUACGAGAACCAGAAGCUGAUCGC CAACCAGUUAACAGCGCCAUCGGCAAGAUGC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUG GGCAAGCUGCAGGACGUGGUGAACCAGAACGC CCAGGCCCUGAACACCCUGGUGAAGCAGCUGA GCAGCAACUUCGGCGCCAUCAGCAGCGUGCUG AACGACAUCCUGAGCCGGCUGGACCCUCCCGA GGCCGAGGUGCAGAUCGACCGGCUGAUCACU GGCCGGCUGCAGAGCCUGCAGACCUACGUGAC CCAGCAGCUGAUCCGGGCGCCGAGAUUCGG GCCAGCGCCAACCUGGCCGCCACCAAGAUGAG CGAGUGCGUGCUGGGCCAGAGCAAGCGGGUG GACUUCUGCGGCAAGGGCUACCACCUGAUGAG CUUUCCCCAGAGCGCACCCCACGGAGUGGUGU UCCUGCACGUGACCUACGUGCCCGCCCAGGAG AAGAACUUCACCACCGCCCCAGCCAUCUGCCA CGACGGCAAGGCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGU GACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAAC UGCGACGUGGUGAUCGGCAUCGUGAACAACAC CGUGUACGAUCCCCUGCAGCCCGAGCUGGACA GCUUCAAGGAGGAGCUGGACAAGUACUUCAAG AAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACA UCCAGAAGGAGAUCGAUCGGCUGAACGAGGUG GCCAAGAACCUGAACGAGAGCCUGAUCGACCU GCAGGAGCUGGGCAAGUACGAGCAGUACAUCA AGUGGCCCUUGGUACAUCUGGCUGGGCUUCAU CGCCGGCCUGAUCGCCAUCGUGAUGGUGACCA	
--	--	--

	UCAUGCUG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLALHRS YLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLVNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKS NLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPI GAGICASYQTQTNSPRRARSVASQSIIAYTMSLGA ENSVAYSNNNSIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDV	20

	VNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPP EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPRE GVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIML	
Cola de PoliA	100 nt	
Variante 7 de la proteína SARS-CoV-2 S		
SEQ ID NO: 21 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 22 y 3' UTR SEQ ID NO: 4.		21
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC	22

AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU
 ACAGCAGCGCCAACAACUGCACCUUCGAGUAC
 GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG
 GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG
 UUCGUGUUCAAGAACAUCGACGGCUACUUCAA
 GAUCUACAGCAAGCACACCCCAAUCAACCUGG
 UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU
 GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA
 ACAUCACCCGGUUCCAGACCCUGCUGGCCCUG
 CACCGGAGCUACCUGACCCCAGGCGACAGCAG
 CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC
 UACGUGGGCUACCUGCAGCCCCGGACCUUCCU
 GCUGAAGUACAACGAGAACGGCACCAUCACCG
 ACGCCGUGGACUGCGCCCUGGACCCUCUGAG
 CGAGACCAAGUGCACCUCUGAAGAGCUUCACCG
 UGGAGAAGGGCAUCUACCAGACCAGCAACUUC
 CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU
 UCCCCAACAUACCAACCUGUGCCCCUUCGGC
 GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU
 GUACGCCUGGAACCGGAAGCGGAUCAGCAACU
 GCGUGGCCGACUACAGCGUGCUGUACAACAGC
 GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU
 GAGCCCCACCAAGCUGAACGACCUGUGCUUCA
 CCAACGUGUACGCCGACAGCUUCGUGAUCCGU
 GGCGACGAGGUGCGGCAGAUCGCACCCGGCC
 AGACAGGCAAGAUCGCCGACUACAACUACAAG
 CUGCCCGACGACUUCACCGGCUGCGUGAUCGC
 CUGGAACAGCAACAACCUCGACAGCAAGGUGG
 GCGGCAACUACAACUACCUGUACCGGCUGUUC
 CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA
 CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA
 CCCCUGCAACGGCGUGGAGGGCUUCAACUGC

	UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGCG UGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCCG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCAUCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCGUGU CACUGAGGAGCGUGGCCAGCCAGAGCAUCAUC GCCUACACCAUGAGCCUGGGCGCCGAGAACAG CGUGGCCUACAGCAACAACAGCAUCGCCAUCC CCACCAACUUCACCAUCAGCGUGACCACCGAG AUUCUGCCCGUGAGCAUGACCAAGACCAGCGU GGACUGCACCAUGUACAUCUGCGGCGACAGCA CCGAGUGCAGCAACCUGCUGCUGCAGUACGGC AGCUUCUGCACCCAGCUGAACCGGGCCCUGAC CGGCAUCGCCGUGGAGCAGGACAAGAACACCC AGGAGGUGUUCGCCCAGGUGAAGCAGAUCUAC AAGACCCCUCCAUCAAGGACUUCGGCGGCUU CAACUUCAGCCAGAUCUGCCCGACCCAGCA AGCCCAGCAAGCGGAGCUUCAUCGAGGACCUG CUGUUCAACAAGGUGACCCUAGCCGACGCCGG	
--	--	--

	CUUCAUCAAGCAGUACGGCGACUGCCUCGGCG ACAUAGCCGCCCCGGGACCUGAUCUGCGCCCAG AAGUUCAACGGCCUGACCGUGCUGCCUCCCCU GCUGACCGACGAGAUGAUCGCCCAGUACACCA GCGCCCUGUUAGCCGGAACCAUCACCAGCGGC UGGACUUUCGGCGCUGGAGCCGCUCUGCAGA UCCCCUUCGCCAUGCAGAUGGCCUACCGGUUC AACGGCAUCGGCGUGACCCAGAACGUGCUGUA CGAGAACCAGAAGCUGAUCGCCAACCAGUUCA ACAGCGCCAUCGGCAAGAUCAGGACAGCCUG AGCAGCACCGCUAGCGCCCUGGGCAAGCUGCA GGACGUGGUGAACCAGAACGCCAGGCCUGA ACACCCUGGUGAAGCAGCUGAGCAGCAACUUC GGCGCCAUCAGCAGCGUGCUGAACGACAUCU GAGCCGGCUGGACAAGGUGGAGGCCGAGGUG CAGAU CGACCGGCUGAUCACUGGCCGGCUGCA GAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCGCGCGAGAUUCGGGCCAGCGCCAA CCUGGCCGCCACCAAGAUGAGCGAGUGCGUGC UGGGCCAGAGCAAGCGGGUGGACUUCUGCGG CAAGGGCUACCACCUGAUGAGCUUUCGCCAGA GCGCACCCACGGAGUGGUGUUCUGCACGU GACCUACGUGCCCGCCCAGGAGAAGAACUUCA CCACCGCCCCAGCCAUCUGCCACGACGGCAAG GCCACUUUCCCCGGGAGGGCGUGUUCGUGA GCAACGGCACCCACUGGUUCGUGACCCAGCGG AACUUCUACGAGCCCCAGAUCAUCACCACCGA CAACACCUUCGUGAGCGGCAACUGCGACGUGG UGAUCGGCAUCGUGAACAACACCGUGUACGAU CCCCUGCAGCCCGAGCUGGACAGCUUCAAGGA GGAGCUGGACAAGUACUUCAAGAAUCACACCA GCCCCGACGUGGACCUGGGCGACAUCAGCGG	
--	--	--

	CAUCAACGCCAGCGUGGUGAACAUCCAGAAGG AGAUCGAUCGGCUGAACGAGGUGGCCAAGAAC CUGAACGAGAGCCUGAUCGACCUGCAGGAGCU GGGCAAGUACGAGCAGUACAUCAAGUGGCCCCU GGUACAUCUGGCUGGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUGU GCUGCAUGACCAGCUGCUGCAGCUGCCUGAAG GGCUGUUGCAGCUGCGGCAGCUGCUGCAAGU UCGACGAGGACGACAGCGAGCCCGUGCUGAAG GGCGUGAAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUUCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFWHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSP TKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLF RKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP	23

	TWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPI GAGICASYQTQTVSLRSVASQSIIAYTMSLGAENS VAYSNNSIAIPTNFTISVTTEILPVSMKTSVDCTMY ICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKP SKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARD LICAQKFNGLTVLPPLLTDemiaQYTSALLAGTITSG WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYEN QKLIANQFNsaIGKIQDSLSTASALGKLQDVVNQN AqALNTLVKQLSSNFGAISSVLNDILSRDKVEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK MSECVLGQSKRVDfCGKGYHLMSFPQSAPHGVV FLHVTYVPAQEKNFTTAPAICHdGKAHFPREGVfV SNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGI VNNTVYDPLQPELDSFKEELDKYfKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNEslIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSC CSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	
Cola de PoliA	100 nt	
Variante 8 de la proteína del SARS-CoV-2 S		
La SEQ ID NO: 24 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 25 y 3' UTR SEQ ID NO: 4.		24
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU	25

codón de detención)	UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGGUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU	
----------------------------	--	--

	GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCUUGCAACGGCGUGGAGGGCUUCAAACUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AAUUC AACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCCG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAAC AGCUACGAGUGCGACAUCCCCAU CGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCAAUU CACCCCGGAGGGCAAGGAGCGUGGCCAGCCA GAGCAUCAUCGCCUACACCAUGAGCCUGGGCG CCGAGAACAGCGUGGCCUACAGCAACAACAGC AUCGCCAUCCCCACCAACUUCACCAUCAGCGU	
--	---	--

	GACCACCGAGAUUCUGCCCCGUGAGCAUGACCA AGACCAGCGUGGACUGCACCAUGUACAUCUGC GGCGACAGCACCGAGUGCAGCAACCUGCUGCU GCAGUACGGCAGCUUCUGCACCCAGCUGAACC GGGCCCUGACCGGCAUCGCCGUGGAGCAGGA CAAGAACACCCAGGAGGUGUUCGCCCAGGUGA AGCAGAUUCUACAAGACCCCUCCAUCAAGGAC UUCGGCGGCUUCAACUUCAGCCAGAUCUGCC CGACCCCAGCAAGCCCAGCAAGCGGAGCUUCA UCGAGGACCUGCUGUUCAACAAGGUGACCCUA GCCGACGCCGGCUUCAUCAAGCAGUACGGCGA CUGCCUCGGCGACAUAAGCCGCCCGGGACCUGA UCUGCGCCCAGAAGUUCAACGGCCUGACCGUG CUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCA UCACCAGCGGCUGGACUUUCGGCGCUGGAGC CGCUCUGCAGAUCCCCUUCGCCAUGCAGAUGG CCUACCGGUUCAACGGCAUCGGCGUGACCCAG AACGUGCUGUACGAGAACCAGAAGCUGAUCGC CAACCAGUUCAACAGCGCCAUCGGCAAGAUC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUG GGCAAGCUGCAGGACGUGGUGAACCAGAACGC CCAGGCCUGAACACCCUGGUGAAGCAGCUGA GCAGCAACUUCGGCGCCAUCAGCAGCGUGCUG AACGACAUCCUGAGCCGGCUGGACAAGGUGGA GGCCGAGGUGCAGAUCGACCGGCUGAUCACU GGCCGGCUGCAGAGCCUGCAGACCUACGUGAC CCAGCAGCUGAUCCGGGCCCGCGAGAUUCGG GCCAGCGCCAACCUGGCCGCCACCAAGAUGAG CGAGUGCGUGCUGGGCCAGAGCAAGCGGGUG GACUUCUGCGGCAAGGGCUACCACCUGAUGAG CUUUCCCCAGAGCGCACCCACGGAGUGGUGU	
--	--	--

	UCCUGCACGUGACCUACGUGCCCCGCCAGGAG AAGAACUUCACCACCGCCCCAGCCAUCUGCCA CGACGGCAAGGCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGU GACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAAC UGCGACGUGGUGAUCGGCAUCGUGAACAACAC CGUGUACGAUCCCCUGCAGCCCGAGCUGGACA GCUUCAAGGAGGAGCUGGACAAGUACUUCAAG AAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACA UCCAGAAGGAGAUCAUCGCGCUGAACGAGGUG GCCAAGAACCUGAACGAGAGCCUGAUCGACCU GCAGGAGCUGGGCAAGUACGAGCAGUACAUCA AGUGGCCCUUGUACAUCUGGCUGGGCUUCAU CGCCGGCCUGAUCGCCAUCGUGAUGGUGACCA UCAUGCUGUGCUGCAUGACCAGCUGCUGCAGC UGCCUGAAGGGCUGUUGCAGCUGCGGCAGCU GCUGCAAGUUCGACGAGGACGACAGCGAGCCC GUGCUGAAGGGCGUG	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE	26

	NGTITDAVDCALDPLSETKCTLSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSP TKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKS NLKP FERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPI GAGICASYQTQTNSPRRARSVASQSIIAYTMSLGA ENSVAYSNN SIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTEC SNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKR SFIEDLLFNKVT LADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPPLL TDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFN SAIGKIQDSL SSTA SALGKLQDV VNQNAQALNTLVKQLSSNFGA ISSVLNDILSR LDKV EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICH DGKAHFPRE GV FVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIA GLIAIVMVTIMLCC MTSCCCLKGCCSCGSCCKFDEDDSEPV LKGV	
Cola de PoliA	100 nt	
Variante 9 de la proteína SARS-CoV-2 S		
La SEQ ID NO: 27 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID		27

NO: 2, ARNm ORF SEQ ID NO: 28 y 3' UTR SEQ ID NO: 4.		
Estructura química	1-metilpseudouridina; SEQ ID NO: 105 corresponde a la secuencia de ARN completamente modificada que incluye la UTR 5', ORF de ARNm y UTR 3'; SEQ ID NO: 106 corresponde al ORF completamente modificado de la construcción de ARNm.	105 106
Cubierta	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA	28

	ACAUCACCCGGUUCCAGACCCUGCUGGCCUG CACCGGAGCUACCUGACCCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUCACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAU CGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCUUGCAACGGCGUGGAGGGCUUCAAUCGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGCCUUAACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC	
--	--	--

	GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCGG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCCAUCCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCAAUU CACCCCGGAGGGCAAGGAGCGUGGCCAGCCA GAGCAUCAUCGCCUACACCAUGAGCCUGGGCG CCGAGAACAGCGUGGCCUACAGCAACAACAGC AUCGCCAUCCCCACCAACUUCACCAUCAGCGU GACCACCGAGAUUCUGCCCGUGAGCAUGACCA AGACCAGCGUGGACUGCACCAUGUACAUCUGC GGCGACAGCACCGAGUGCAGCAACCUGCUGCU GCAGUACGGCAGCUUCUGCACCCAGCUGAACC GGGCCUGACCGGCAUCGCCGUGGAGCAGGA CAAGAACACCCAGGAGGUGUUCGCCCAGGUGA AGCAGAUCAACAAGACCCCUCCCAUCAAGGAC UUCGGCGGCUUCAACUUCAGCCAGAUCUGCC CGACCCAGCAAGCCCAGCAAGCGGAGCUUCA UCGAGGACCUGCUGUUCAACAAGGUGACCCUA GCCGACGCCGGCUUCAUCAAGCAGUACGGCGA CUGCCUCGGCGACAUAAGCCGCCCGGGACCUGA UCUGCGCCCAGAAGUUCAACGGCCUGACCGUG CUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCUGUUAAGCCGGAACCA UCACCAGCGGCUGGACUUUCGGCGCUGGAGC CGCUCUGCAGAUCCCCUUCGCCAUGCAGAUGG CCUACCGGUUCAACGGCAUCGGCGUGACCCAG	
--	--	--

	AACGUGCUGUACGAGAACCAGAAGCUGAUCGC CAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUG GGCAAGCUGCAGGACGUGGUGAACCAAGAACGC CCAGGCCCUGAACACCCUGGUGAAGCAGCUGA GCAGCAACUUCGGCGCCAUCAGCAGCGUGCUG AACGACAUCCUGAGCCGGCUGGACCCUCCCGA GGCCGAGGUGCAGAUUCGACCGGCUGAUCACU GGCCGGCUGCAGAGCCUGCAGACCUACGUGAC CCAGCAGCUGAUCCGGGCGCCGAGAUUCGG GCCAGCGCCAACCUGGCCGCCACCAAGAUGAG CGAGUGCGUGCUGGGCCAGAGCAAGCGGGUG GACUUCUGCGGCAAGGGCUACCACCUGAUGAG CUUUCCCCAGAGCGCACCCCACGGAGUGGUGU UCCUGCACGUGACCUACGUGCCCGCCCAGGAG AAGAACUUCACCACCGCCCCAGCCAUCUGCCA CGACGGCAAGGCCACUUCUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGU GACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAAC UGCGACGUGGUGAUCGGCAUCGUGAACAACAC CGUGUACGAUCCCCUGCAGCCCGAGCUGGACA GCUUCAAGGAGGAGCUGGACAAGUACUUCAAG AAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACA UCCAGAAGGAGAUUCGAUCGGCUGAACGAGGUG GCCAAGAACCUGAACGAGAGCCUGAUCGACCU GCAGGAGCUGGGCAAGUACGAGCAGUACAUCA AGUGGCCCUUGGUACAUCUGGCUGGGCUUCAU CGCCGGCCUGAUCGCCAUCGUGAUGGUGACCA UCAUGCUGUGCUGCAUGACCAGCUGCUGCAGC UGCCUGAAGGGCUGUUGCAGCUGCGGCAGCU	
--	---	--

	GCUGCAAGUUCGACGAGGACGACAGCGAGCCC GUGCUGAAGGGCGUGAAGCUGCACUACACC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLF RKS NLKP FERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSECDIPI GAGICASYQTQTNSPRRARSVASQSIIAYTMSLGA ENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV	29

	LYENQKLIANQFNSAIGKIQDSLSTASALGKLQDV VNQNAQALNTLVKQLSSNFGAISSVLNDILSRDPP EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPRE GVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCC MTSCCCLKGCCSCGSCCKFDEDDSEPVKGVKL HYT	
Cola de PoliA	100 nt	
Variante 10 de la proteína SARS-CoV-2 S		
SEQ ID NO: 51 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 52 y 3' UTR SEQ ID NO: 4.		51
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG	52

	CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGGUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUCACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAU CGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC	
--	--	--

	CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCCUUGCAACGGCGUGGAGGGCUUCAAUCGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AAUUCUACACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCGG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCCAUCCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCAAUU CACCCCGGAGGGGCAAGGAGCGUGGCCAGCCA GAGCAUCAUCGCCUACACCAUGAGCCUGGGCG CCGAGAACAGCGUGGCCUACAGCAACAACAGC AUCGCCAUCCCCACCAACUUCACCAUCAGCGU GACCACCGAGAUUCUGCCCGUGAGCAUGACCA AGACCAGCGUGGACUGCACCAUGUACAUCUGC GGCGACAGCACCGAGUGCAGCAACCUGCUGCU GCAGUACGGCAGCUUCUGCACCCAGCUGAACC GGGCCUGACCGGCAUCGCCGUGGAGCAGGA CAAGAACACCCAGGAGGUGUUCGCCCAGGUGA AGCAGAUCAACAAGACCCCUCCCAUCAAGGAC	
--	---	--

	UUCGGCGGCUUCAACUUCAGCCAGA UCCUGCC CGACCCCAGCAAGCCCAGCAAGCGGAGCUUCA UCGAGGACCUGCUGUUCAACAAGGUGACCCUA GCCGACGCCGGCUUCAUCAAGCAGUACGGCGA CUGCCUCGGCGACAUAAGCCGCCCGGGACCUGA UCUGCGCCCAGAAGUUCAACGGCCUGACCGUG CUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCA UCACCAGCGGCUGGACUUUCGGCGCUGGAGC CGCUCUGCAGA UCCCCUUCGCCAUGCAGAUGG CCUACCGGUUCAACGGCAUCGGCGUGACCCAG AACGUGCUGUACGAGAACCAGAAGCUGAUCGC CAACCAGUUCAACAGCGCCAUCGGCAAGAUCC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUG GGCAAGCUGCAGGACGUGGUGAACCAGAACGC CCAGGCCCUGAACACCCUGGUGAAGCAGCUGA GCAGCAACUUCGGCGCCAUCAGCAGCGUGCUG AACGACA UCCUGAGCCGGCUGGACAAGGUGGA GGCCGAGGUGCAGAUCGACCGGCUGAUCACU GGCCGGCUGCAGAGCCUGCAGACCUACGUGAC CCAGCAGCUGA UCCGGGCCGCCGAGAUUCGG GCCAGCGCCAACCUGGCCGCCACCAAGAUGAG CGAGUGCGUGCUGGGCCAGAGCAAGCGGGUG GACUUCUGCGGCAAGGGCUACCACCUGAUGAG CUU UCCCCAGAGCGCACCCACGGAGUGGUGU UCCUGCACGUGACCUACGUGCCCGCCCAGGAG AAGAACUUCACCACCGCCCCAGCCAUCUGCCA CGACGGCAAGGCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGU GACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAAC UGCGACGUGGUGAUCGGCAUCGUGAACAACAC	
--	--	--

	CGUGUACGAUCCCCUGCAGCCCGAGCUGGACA GCUUCAAGGAGGAGCUGGACAAGUACUUCAAG AAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACA UCCAGAAGGAGAUCGAUCGGCUGAACGAGGUG GCCAAGAACCUGAACGAGAGCCUGAUCGACCU GCAGGAGCUGGGCAAGUACGAGCAGUACAUCA AGUGGCCCUGGUACAUCUGGCUGGGCUUCAU CGCCGGCCUGAUCGCCAUCGUGAUGGUGACCA UCAUGCUGUGCUGCAUGACCAGCUGCUGCAGC UGCCUGAAGGGCUGUUGCAGCUGCGGCAGCU GCUGCAAGUUCGACGAGGACGAC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKS NLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL	33

	PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPI GAGICASYQTQTNSPRRARSVASQSIIAYTMSLGA ENSVAYSNNNSIAIPTNFTISVTTEILPVSMKTSVDC TMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKRSEFIEDLLFNKVTLADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFNSAIGKIQDSLSTASALGKLQDV VNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKV EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICHGKAHFPRE GVSVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCC MTSCCSCCLKGCCSCGSCCKFDEDD	
Cola de PoliA	100 nt	
Variante 11 de la proteína SARS-CoV-2 S		
SEQ ID NO: 53 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 54 y 3' UTR SEQ ID NO: 4.		53
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAAUAAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG	54

ARNm (excluyendo el codón de detención)	ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU	
--	---	--

	GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCUUGCAACGGCGUGGAGGGCUUCAAACUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCGG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCAUCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCAAUU CACCCCGGAGGGCAAGGAGCGUGGCCAGCCA GAGCAUCAUCGCCUACACCAUGAGCCUGGGCG	
--	--	--

	CCGAGAACAGCGUGGCCUACAGCAACAACAGC AUCGCCAUCCCCACCAACUUCACCAUCAGCGU GACCACCGAGAUUCUGCCCGUGAGCAUGACCA AGACCAGCGUGGACUGCACCAUGUACAUCUGC GGCGACAGCACCGAGUGCAGCAACCUGCUGCU GCAGUACGGCAGCUUCUGCACCCAGCUGAACC GGGCCUGACCGGCAUCGCCGUGGAGCAGGA CAAGAACACCCAGGAGGUGUUCGCCCAGGUGA AGCAGAUUCUACAAGACCCCUCCCAUCAAGGAC UUCGGCGGCUUCAACUUCAGCCAGAUCUGCC CGACCCAGCAAGCCCAGCAAGCGGAGCUUCA UCGAGGACCUGCUGUUCAACAAGGUGACCCUA GCCGACGCCGGCUUCAUCAAGCAGUACGGCGA CUGCCUCGGCGACAUAAGCCGCCCGGGACCUGA UCUGCGCCCAGAAGUUCAACGGCCUGACCGUG CUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAAGCCGGAACCA UCACCAGCGGCUGGACUUUCGGCGCUGGAGC CGCUCUGCAGAUCCCCUUCGCCAUGCAGAUGG CCUACCGGUUCAACGGCAUCGGCGUGACCCAG AACGUGCUGUACGAGAACCAGAAGCUGAUCGC CAACCAGUUCAACAGCGCCAUCGGCAAGAUGC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUG GGCAAGCUGCAGGACGUGGUGAACCAGAACGC CCAGGCCCUGAACACCCUGGUGAAGCAGCUGA GCAGCAACUUCGGCGCCAUCAGCAGCGUGCUG AACGACAUCCUGAGCCGGCUGGACCCUCCCGA GGCCGAGGUGCAGAUUCGACCGGCUGAUCACU GGCCGGCUGCAGAGCCUGCAGACCUACGUGAC CCAGCAGCUGAUCCGGGCCCGCGAGAUUCGG GCCAGCGCCAACCUGGCCGCCACCAAGAUGAG CGAGUGCGUGCUGGGCCAGAGCAAGCGGGUG	
--	--	--

	GACUUCUGCGGCAAGGGCUACCACCUGAUGAG CUUUCCCCAGAGCGCACCCCACGGAGUGGUGU UCCUGCACGUGACCUACGUGCCCGCCCAGGAG AAGAACUUCACCACCGCCCCAGCCAUCUGCCA CGACGGCAAGGCCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGU GACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAAC UGCGACGUGGUGAUCGGCAUCGUGAACAACAC CGUGUACGAUCCCCUGCAGCCCGAGCUGGACA GCUUCAAGGAGGAGCUGGACAAGUACUUCAAG AAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACA UCCAGAAGGAGAUCGAUCGGCUGAACGAGGUG GCCAAGAACCUGAACGAGAGCCUGAUCGACCU GCAGGAGCUGGGCAAGUACGAGCAGUACAUCA AGUGGCCCUGGUACAUCUGGCUGGGCUUCAU CGCCGGCCUGAUCGCCAUCGUGAUGGUGACCA UCAUGCUGUGCUGCAUGACCAGCUGCUGCAGC UGCCUGAAGGGCUGUUGCAGCUGCGGCAGCU GCUGCAAGUUCGACGAGGACGAC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUUCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS	34

	YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSP TKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSECDIPI GAGICASYQTQTNSPRRARSVASQSIIAYTMSLGA ENSVAYSNNNSIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTEC SNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKRSFIEDLLFNKVT LADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLP LLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFNSAIGKIQDSLSTASALGKLQDV VNQNAQALNTLVKQLSSNFGAISSVLNDILSR LDPP EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICH DGKAHFPRE GV FVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCC MTSCC SCLKGCCSCGSCCKFDEDD	
Cola de PoliA	100 nt	
Variante 12 de la proteína del SARS-CoV-2 S		

La SEQ ID NO: 55 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 56 y 3' UTR SEQ ID NO: 4.		55
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	UUUCAACGACGGCGUGUACUUCGCCAGCACCG AGAAGAGCAACAUCAUCCGGGGCUGGAUCUUC GGCACCACCCUGGACAGCAAGACCCAGAGCCU GCUGAUCGUGAAUAACGCCACCAACGUGGUGA UCAAGGUGUGCGAGUUCCAGUUCUGCAACGAC CCCUUCCUGGGCGUGUACUACCACAAGAACAA CAAGAGCUGGAUGGAGAGCGAGUUCCGGGUG UACAGCAGCGCCAACAACUGCACCUUCGAGUA CGUGAGCCAGCCCUUCCUGAUGGACCUGGAG GGCAAGCAGGGCAACUUCAAGAACCUGCGGGA GUUCGUGUUCAAGAACAUCGACGGCUACUUCA AGAUCUACAGCAAGCACACCCCAAUCAACCUG GUGCGGGAUCUGCCCCAGGGCUUCUCAGCCC UGGAGCCCCUGGUGGACCUGCCCAUCGGCAU CAACAUCACCCGGUUCCAGACCCUGCUGGCCC UGCACCGGAGCUACCUGACCCCAGGCGACAGC AGCAGCGGGUGGACAGCAGGCGCGGCUGCUU ACUACGUGGGCUACCUGCAGCCCCGGACCUUC CUGCUGAAGUACAACGAGAACGGCACCAUCAC CGACGCCGUGGACUGCGCCCUGGACCCUCUG AGCGAGACCAAGUGCACCCUGAAGAGCUUCAC CGUGGAGAAGGGCAUCUACCAGACCAGCAACU UCCGGGUGCAGCCCACCGAGAGCAUCGUGCG GUUCCCCAACAUACCAACCUGUGCCCCUUCG	56

	GCGAGGUGUUCAACGCCACCCGGUUCGCCAGC GUGUACGCCUGGAACCGGAAGCGGAUCAGCAA CUGCGUGGCCGACUACAGCGUGCUGUACAACA GCGCCAGCUUCAGCACCUUCAAGUGCUACGGC GUGAGCCCCACCAAGCUGAACGACCUGUGCUU CACCAACGUGUACGCCGACAGCUUCGUGAUCC GUGGCGACGAGGUGCGGCAGAUCGCACCCGG CCAGACAGGCAAGAUCGCCGACUACAACUACA AGCUGCCCGACGACUUCACCGGCUGCGUGAUC GCCUGGAACAGCAACAACCUCGACAGCAAGGU GGGCGGCAACUACAACUACCUGUACCGGCUGU UCCGGAAGAGCAACCUGAAGCCCUUCGAGCGG GACAUCAGCACCGAGAUCUACCAAGCCGGCUC CACCCCUUGCAACGGCGUGGAGGGCUUCAACU GCUACUUCCCUUCUGCAGAGCUACGGCUUCCAG CCCACCAACGGCGUGGGCUACCAGCCCUACCG GGUGGUGGUGCUGAGCUUCGAGCUGCUGCAC GCCCCAGCCACCGUGUGUGGCCCCAAGAAGAG CACCAACCUGGUGAAGAACAAGUGCGUGAACU UCAACUUAACGGCCUUAACCGGCACCGGCGUG CUGACCGAGAGCAACAAGAAAUUCCUGCCCUU UCAGCAGUUCGGCCGGGACAUCGCCGACACCA CCGACGCUGUGCGGGAUCCCCAGACCCUGGA GAUCCUGGACAUCACCCCUUGCAGCUUCGGCG GCGUGAGCGUGAUCACCCAGGCACCAACACC AGCAACCAGGUGGCCGUGCUGUACCAGGACGU GAACUGCACCGAGGUGCCCGUGGCCAUCCACG CCGACCAGCUGACACCCACCUGGCGGGUCUAC AGCACCGGCAGCAACGUGUUCAGACCCGGGC CGGUUGCCUGAUCGGCGCCGAGCACGUGAACA ACAGCUACGAGUGCGACAUCCCCAUCCGGCGCC GGCAUCUGUGCCAGCUACCAGACCCAGACCAA	
--	---	--

	UUCACCCGGCAGCGGCGGCAGCGUGGCCAGC CAGAGCAUCAUCGCCUACACCAUGAGCCUGGG CGCCGAGAACAGCGUGGCCUACAGCAACAACA GCAUCGCCAUCCCCACCAACUUCACCAUCAGC GUGACCACCGAGAUUCUGCCCGUGAGCAUGAC CAAGACCAGCGUGGACUGCACCAUGUACAUCU GCGGCGACAGCACCGAGUGCAGCAACCUGCUG CUGCAGUACGGCAGCUUCUGCACCCAGCUGAA CCGGGCCCUGACCGGCAUCGCCGUGGAGCAG GACAAGAACACCCAGGAGGUGUUCGCCCAGGU GAAGCAGAUUCUACAAGACCCCUCCCAUCAAGG ACUUCGGCGGCUUCAACUUCAGCCAGAUCUG CCCGACCCAGCAAGCCCAGCAAGCGGAGCUU CAUCGAGGACCUGCUGUUCAACAAGGUGACCC UAGCCGACGCCGGCUUCAUCAAGCAGUACGGC GACUGCCUCGGCGACAUAGCCGCCCGGGACCU GAUCUGCGCCCAGAAGUUCAACGGCCUGACCG UGCUGCCUCCCCUGCUGACCGACGAGAUGAUC GCCCAGUACACCAGCGCCCUGUUAAGCCGGAAC CAUCACCAGCGGCUGGACUUUCGGCGCUGGA GCCGCUCUGCAGAUCCCCUUCGCCAUGCAGAU GGCCUACCGGUUCAACGGCAUCGGCGUGACCC AGAACGUGCUGUACGAGAACCAGAAGCUGAUC GCCAACCAGUUCAACAGCGCCAUCGGCAAGAU CCAGGACAGCCUGAGCAGCACCGCUAGCGCCC UGGGCAAGCUGCAGGACGUGGUGAACCAGAAC GCCAGGCCCUGAACACCCUGGUGAAGCAGCU GAGCAGCAACUUCGGCGCCAUCAGCAGCGUGC UGAACGACAUCUGAGCCGGCUGGACCCUCCC GAGGCCGAGGUGCAGAUCGACCGGCUGAUCAC UGGCCGGCUGCAGAGCCUGCAGACCUACGUGA CCCAGCAGCUGAUCCGGGCGCGCGAGAUUCG	
--	--	--

	GGCCAGCGCCAACCUGGCCGCCACCAAGAUGA GCGAGUGCGUGCUGGGGCCAGAGCAAGCGGGU GGACUUCUGCGGCAAGGGCUACCACCUGAUGA GCUUUCCCCAGAGCGCACCCACGGAGUGGUG UUCCUGCACGUGACCUACGUGCCCGCCCAGGA GAAGAAUUCACCACCGCCCCAGCCAUCUGCC ACGACGGCAAGGCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGU GACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAAC UGCGACGUGGUGAUCGGCAUCGUGAACAACAC CGUGUACGAUCCCCUGCAGCCCGAGCUGGACA GCUUCAAGGAGGAGCUGGACAAGUACUUCAAG AAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACA UCCAGAAGGAGAUCGAUCGGCUGAACGAGGUG GCCAAGAACCUGAACGAGAGCCUGAUCGACCU GCAGGAGCUGGGCAAGUACGAGCAGUACAUCA AGUGGCCCUUGUACAUCUGGCUGGGCUUCAU CGCCGGCCUGAUCGCCAUCGUGAUGGUGACCA UCAUGCUGUGCUGCAUGACCAGCUGCUGCAGC UGCCUGAAGGGCUGUUGCAGCUGCGGCAGCU GCUGCAAGUUCGACGAGGACGAC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP	35

	FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLF RKS NLKPFERDISTEIQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQP YRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSECDIPI GAGICASYQTQTNSPGSGGSVASQSIIAYTMSLGA ENSVAYSNNNSIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKR SFIEDLLFNKVTLADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDV VNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPP EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICHGKAHFPRE GVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCC MTSCCSCCLKGCCSCGSCCKFDEDD	
Cola de PoliA	100 nt	

Variante 1 de proteína WIV16 S		
La SEQ ID NO: 57 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 48 y 3' UTR SEQ ID NO: 4.		57
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NImpNp	
5' UTR	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUUAUCUUCCUGUUCUUCCUGACCCUGAC CAGCGGCAGCGACCUGGAAAGCUGCACCACCU UCGACGACGUGCAGGCCCCCAACUACCCUCAG CACAGCUCUAGCAGACGGGGCGUGUACUACCC CGACGAGAUUCUUCAGAAGCGACACCCUGUACC UGACCCAGGACCUGUUCCUGCCCUUCUACAGC AACGUGACCGGCUUCCACACCAUCAACCACAG AUUCGACAACCCCGUGAUCCCCUUCAAGGACG GGGUGUACUUUGCCGCCACCGAGAAGUCCAAU GUCGUGCGGGGAUGGGUGUUCGGCAGCACCA UGAACAACAAGAGCCAGAGCGUGAUCAUCAUC AACAAACAGCACCAACGUCGUGAUCCGGGCCUG CAACUUCGAGCUGUGCGACAACCCAUUCUUCG CCGUGUCCAAGCCCACCGGCACCCAGACCCAC ACCAUGAUCUUCGACAACGCCUUAACUGCAC CUUCGAGUACAUCAGCGACAGCUUCAGCCUGG ACGUGGCCGAGAAAAGCGGCAACUUCAAGCAC CUGAGAGAAUUCGUGUUCAAGAACAAGGACGG CUUCCUGUACGUGUACAAGGGCUACCAGCCCA UCGACGUCGUGCGGAUCUGCCCAGCGGCUU CAACAUCCUGAAGCCCAUCUUCAAGCUGCCCC UGGGCAUCAACAUCACCAACUUCGGGGCUAUC CUGACCGCCUUCUGCCCGCCCAGGAUACCUG	48

	GGGAACAAGCGCCGCUGCCUACUUCGUGGGC UACCUGAAGCCUGCCACCUUCAUGCUGAAGUA CGACGAGAACGGCACCAUCACCGACGCCGUGG ACUGCAGCCAGAAUCCUCUGGCCGAGCUGAAG UGCAGCGUGAAGUCCUUCGAGAUCGACAAGGG CAUCUACCAGACCAGCAACUUCAGAGUGGCCC CCAGCAAAGAAGUCGUGCGGUUCCCCAAUAUC ACCAACCUGUGCCCCUUCGGCGAGGUGUUCAA CGCCACCACCUUUCAGCGUGUACGCCUGGG AGCGGAAGCGGAUCAGCAACUGCGUGGCCGAC UACAGCGUGCUGUACAACUCCACCAGCUUCUC CACCUUCAAGUGCUACGGCGUGUCCGCCACCA AGCUGAACGACCUGUGCUUCAGCAAUGUGUAC GCCGACUCCUUCGUCGUGAAGGGCGACGAUG UGCGCCAGAUCCGCCUUGGACAGACAGGCGUG AUCGCCGAUUACAACUACAAGCUGCCUGACGA CUUCACCGGCUGCGUGCUGGCCUGGAACACCA GAAACAUCGACGCCACCCAGACAGGCAACUAC AAUUACAAGUACAGAAGCCUGCGGCACGGCAA GCUGCGGCCCUCGAGAGGGACAUCUCCAACG UGCCCUUCAGCCCCGACGGCAAGCCUUGUACC CCCCUGCCUUUAACUGCUACUGGCCCCUGAA CGACUACGGCUUCUACAUCACAAACGGCAUCG GCUAUCAGCCCUACCGGGUGGUGGUGCUGUC CUUUGAGCUGCUGAAUGCCCCUGCCACCGUGU GCGGCCCUAAGCUGAGCACCGACCUGAUCAAG AACCAGUGCGUGAACUUAACUUAACGGCCU GACCGGCACCGGCGUGCUGACACCUAGCAGCA AGAGAUUCCAGCCCUUCCAGCAGUUCGGCCGG GACGUGCUGGAUUUCACCGACAGCGUGCGGG ACCCAAGACCAGCGAGAUCCUGGACAUCAGC CCCUGCAGCUUCGGCGGAGUGUCCGUGAUCA	
--	--	--

	CCCCCGGCACCAAUACCAGCUCUGAGGUGGCC GUGCUGUAUCAGGACGUGAACUGCACCGAUGU GCCCGUGGCCAUCCACGCCGAUCAGCUGACCC CAUCUUGGCGGGUGUACUCCACCGGCAACAAC GUGUUCCAGACACAAGCCGGCUGCCUGAUCGG AGCCGAGCACGUGGACACCAGCUACGAGUGCG ACAUCCCUAUCGGCGCUGGCAUCUGCGCCAGC UACCACACCGUGUCCAGCCUGAGAAGCACCAG CCAGAAAUCUAUCGUGGCCUACACCAUGAGCC UGGGCGCCGACAGCUCUAUCGCCUACUCCAAC AACACAAUCGCCAUCCCCACCAAUUUCAGCAUC UCCAUCACCACCGAAGUGAUGCCCGUGUCCAU GGCCAAGACCUCGUGGAUUGCAACAUGUACA UCUGCGGCGACAGCACCGAGUGCGCCAACCUG CUGCUGCAGUACGGCAGCUUCUGCACCCAGCU GAACAGAGCCCUGAGCGGAAUCGCCGUGGAAC AGGACAGAAACACCCGGGAAGUGUUCGCCCAA GUGAAGCAGAUGUUAUAAGACCCCCACCCUGAA GGAUUUCGGCGGCUUUAACUUCAGCCAGAUC UGCCCGACCCUCUGAAGCCUACCAAGCGGAGC UUCAUCGAGGACCUGCUGUUAACAAAGUGAC CCUGGCCGACGCCGGCUUUAUGAAGCAGUAUG GCGAGUGCCUGGGCGACAUCAACGCCCGGGA UCUGAUCUGCGCCCAGAAGUUUAACGGACUGA CCGUGCUGCCCCCUCUGCUGACCGACGAUAUG AUCGCCGCCUACACAGCCGCCUGGUGUCUGG CACAGCUACCGCCGGAUGGACAUUUGGAGCUG GCGCCGCUCUGCAGAUCCCCUUGCCAUGCAG AUGGCCUACCGGUUCAUUGGCAUCGGCGUGAC CCAGAAUGUGCUGUACGAGAACCAGAAGCAGA UCGCCAACCAGUUCAACAAGGCCAUUAGCCAG AUUCAGGAAAGCCUGACCACCACCAGCACCGC	
--	--	--

	CCUGGGCAAACUGCAGGACGUCGUGAACCAGA ACGCCCAGGCCCUGAACACCCUCGUGAAGCAG CUGAGCAGCAAUUUCGGCGCCAUCAGCUCCGU GCUGAACGAUAUCCUGAGCAGACUGGACAAGG UGGAAGCAGAGGUGCAGAUCCGACCGGCUGAUC ACCGGCAGACUGCAGAGCCUGCAGACCUACGU GACACAGCAGCUGAUUAGAGCCGCCGAGAUA GGGCCAGCGCCAAUCUGGCCGCCACAAAGAUG AGCGAGUGUGUGCUGGGCCAGAGCAAGCGGG UGGACUUCUGCGGCAAGGGCUAUCACCUGAUG AGCUUCCCCCAGGCCGCUCCUCACGGCGUGG UGUUUCUGCACGUGACAUACGUGCCCAGCCAG GAACGGAACUUCACCACCGCCCCAGCCAUCUG CCACGAGGGCAAGGCCUACUUCCCCCGGGAAG GCGUGUUCGUGUUUAACGGCACCUCUGGUU UAUCACCCAGCGGAUUUCUUCAGUCCGCAGA UCAUCACCACAGACAACACCUUCGUGUCCGGC AGCUGCGACGUCGUGAUUGGCAUCAUUAACAA CACCGUGUACGACCCCCUGCAGCCCGAGCUGG ACAGCUUCAAGAGGAACUGGACAAGUACUUC AAGAACCACACCUCCCCCGACGUGGACCUGGG CGAUUUCUCCGGCAUCAUUGCCAGCGUCGUGA AUAUCCAGAAAGAGAUCAUCGCCUGAACGAG GUGGCCAAGAACCUGAAUGAGAGCCUGAUCGA CCUGCAGGAACUGGGGAAGUACGAGCAGUACA UCAAGUGGCCUUGGUACGUGUGGCUGGGCUU UAUCGCCGGCCUGAUCGCCAUCGUGAUGGUCA CCAUCCUGCUGUGCUGCAUGACCAGCUGUUGC AGCUGUCUGAAGGGCGCCUGCAGCUGUGGCU CCUGCUGCAAGUUCGAUGAGGACGACAGCGAG CCUGUGCUGAAAGGCGUGAAGCUGCACUACAC C	
--	---	--

3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUCCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFIFLFFLTLTSGSDLESCCTTFDDVQAPNYPQHSS SRRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFH TINHRFDNPVIPFKDGVYFAATEKSNVVRGWVFGS TMNKSQSVIIINNSTNVVIRACNFELCDNPFFAVS KPTGTQTHTMIFDNAFNCTFEYISDSFSLDVAEKS GNFKHLREFVFKNKDGLYVYKGYQPIDVVRDLPS GFNILKPIFKLPLGINITNFRAILTAFLPAQDTWGTSA AAYFVGYLKPATFMLKYDENGITITDAVDCSQNP ELKCSVKSFEIDKGIYQTSNFRVAPSKEVVRFPNIT NLCPFGEVFNATTFPSVYAWERKRISNCVADYSVL YNSTSFSSTFKCYGVSATKLNDLCFSNVYADSFVVK GDDVRQIAPGQTGVIADYNYKLPDDFTGCVLAWN TRNIDATQTGNVNYKYRSLRHGKLRPFERDISNVP FSPDGKPCTPPAFNCYWPLNDYGFYITNGIGYQPY RVVLSFELLNAPATVCGPKLSTDLIKNCVNFNF NGLTGTGVLTPSSKRFQPFQQFGRDVLDFDTSVR DPKTSEILDISPCSFGGVSVITPGTNTSSEVAVLYQ DVNCTDVPVAIHADQLTPSWRVYSTGNNVFQTQA GCLIGAEHVDTSYECDIPIGAGICASYHTVSSLRST SQKSIVAYTMSLGADSSIAYSNNITIAIPTNFSISITTE VMPVSMAKTSVDCNMYICGDSTECANLLLQYGSF CTQLNRALSGIAVEQDRNTREVFAQVKQMYKTPT LKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLA DAGFMKQYGECLGDINARDLICAQKFNGLTVLPPL LTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFA MQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQI QESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSN FGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQT	47

	YVTQQLIRAAEIRASANLAATKMSECVLGQSKRVD FCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNF TTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFF SPQIITTDNTFVSGSCDVVIGIINNTVYDPLQPELDS FKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEI DRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWL GFIAGLIAIVMTILLCCMTSCCSCCLKGACSCGSCC KFDEDDSEPVLKGVKLHYT	
Cola de PoliA	100 nt	
Variante 2 de proteína WIV16 S		
La SEQ ID NO: 58 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 50 y 3' UTR SEQ ID NO: 4.		58
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCAUCUCCUGUUCUCCUGACCCUGAC CAGCGGCAGCGACCUGGAGAGCUGCACCACCU UCGACGACGUGCAGGCCCCUAACUACCCUCAG CACAGCAGCAGCAGAAGAGGCGUGUACUACCC UGACGAGAUCUUCAGAAGCGACACCCUGUACC UGACCCAGGACCUGUCCUGCCUUUCUACAGC AACGUGACCGGCUUCCACACCAUCAACCACAG AUUCGACAACCCUGUGAUCCCUUUAAGGACG GCGUGUACUUCGCCGCCACCGAGAAGAGCAAC GUGGUGAGAGGCUGGGUGUUCGGCAGCACCA UGAACAAACAAGAGCCAGAGCGUGAUCAUCAUC AACAAACAGCACCAACGUGGUGAUCAGAGCCUG CAACUUCGAGCUGUGCGACAACCCUUUCUUCG CCGUGAGCAAGCCUACCGGCACCCAGACCCAC	50

ACCAUGAUCUUCGACAACGCCUUAACUGCAC
 CUUCGAGUACAUCAGCGACAGCUUCAGCCUGG
 ACGUGGCCGAGAAGAGCGGCAACUUCAAGCAC
 CUGAGAGAGUUCGUGUUCAAGAACAAGGACGG
 CUUCCUGUACGUGUACAAGGGCUACCAGCCUA
 UCGACGUGGUGAGAGACCUGCCUAGCGGCUU
 CAACAUCCUGAAGCCUAUCUUCAAGCUGCCUC
 UGGGCAUCAACAUCACCAACUUCAGAGCCAUC
 CUGACCGCCUUCUGCCUGCCCAGGACACCUG
 GGGCACCAGCGCCGCCGCCUACUUCGUGGGC
 UACCUGAAGCCUGCCACCUUCAUGCUGAAGUA
 CGACGAGAACGGCACCAUCACCGACGCCGUGG
 ACUGCAGCCAGAACCCUCUGGCCGAGCUGAAG
 UGCAGCGUGAAGAGCUUCGAGAUCGACAAGGG
 CAUCUACCAGACCAGCAACUUCAGAGUGGCCC
 CUAGCAAGGAGGUGGUGAGAUUCCCUAACAUC
 ACCAACCUGUGCCCUUUCGGCGAGGUGUUCAA
 CGCCACCACCUUCCCUAGCGUGUACGCCUGGG
 AGAGAAAGAGAAUCAGCAACUGCGUGGCCGAC
 UACAGCGUGCUGUACAACAGCACCAAGCUUCAG
 CACCUUCAAGUGCUACGGCGUGAGCGCCACCA
 AGCUGAACGACCUGUGCUUCAGCAACGUGUAC
 GCCGACAGCUUCGUGGUGAAGGGCGACGACG
 UGAGACAGAUCGCCCCUGGCCAGACCGGGCGUG
 AUCGCCGACUACAACUACAAGCUGCCUGACGA
 CUUCACCGGCUGCGUGCUGGCCUGGAACACCA
 GAAACAUCGACGCCACCCAGACCGGCAACUAC
 AACUACAAGUACAGAAGCCUGAGACACGGCAA
 GCUGAGACCUUUCGAGAGAGACAUCAGCAACG
 UGCCUUUCAGCCCUGACGGCAAGCCUUGCACC
 CCUCCUGCCUUAACUGCUACUGGCCUCUGAA
 CGACUACGGCUUCUACAUCACCAACGGCAUCG

	GCUACCAGCCUUACAGAGUGGUGGUGCUGAGC UUCGAGCUGCUGAACGCCCCUGCCACCGUGUG CGGCCCUAAGCUGAGCACCGACCUGAUCAAGA ACCAGUGCGUGAACUUCAACUUCAACGGCCUG ACCGGCACCGGCGUGCUGACCCCUAGCAGCAA GAGAUUCCAGCCUUUCCAGCAGUUCGGCAGAG ACGUGCUGGACUUCACCGACAGCGUGAGAGAC CCUAAGACCAGCGAGAUCUGGACAUCAGCCC UUGCAGCUUCGGCGGCGUGAGCGUGAUCACC CCUGGCACCAACACCAGCAGCGAGGUGGCCGU GCUGUACCAGGACGUGAACUGCACCGACGUGC CUGUGGCCAUCCACGCCGACCAGCUGACCCCU AGCUGGAGAGUGUACAGCACCGGCAACAACGU GUUCCAGACCCAGGCCGGCUGCCUGAUCGGC GCCGAGCACGUGGACACCAGCUACGAGUGCGA CAUCCCUAUCGGCGCCGGCAUCUGCGCCAGCU ACCACACCGUGAGCAGCCUGAGAAGCACCCAGC CAGAAGAGCAUCGUGGCCUACACCAUGAGCCU GGGCGCCGACAGCAGCAUCGCCUACAGCAACA ACACCAUCGCCAUCCCUACCAACUUCAGCAUCA GCAUCACCACCGAGGUGAUGCCUGUGAGCAUG GCCAAGACCAGCGUGGACUGCAACAUGUACAU CUGCGGCGACAGCACCGAGUGCGCCAACCUGC UGCUGCAGUACGGCAGCUUCUGCACCCAGCUG AACAGAGCCCUGAGCGGCAUCGCCGUGGAGCA GGACAGAAACACCAGAGAGGUGUUCGCCCAGG UGAAGCAGAUGUACAAGACCCCUACCCUGAAG GACUUCGGCGGCUUCAACUUCAGCCAGAUCU GCCUGACCCUCUGAAGCCUACCAAGAGAAGCU UCAUCGAGGACCUGCUGUUCAACAAGGUGACC CUGGCCGACGCCGGCUUCAUGAAGCAGUACGG CGAGUGCCUGGGCGACAUCAACGCCAGAGACC	
--	---	--

	UGAUCUGCGCCCAGAAGUUCAACGGCCUGACC GUGCUGCCUCCUCUGCUGACCGACGACAUGAU CGCCGCCUACACCGCCGCCUUGGUGAGCGGC ACCGCCACCGCCGGCUGGACCUUCGGCGCCG GCGCCGCCUUGCAGAUCCCUUUCGCCAUGCAG AUGGCCUACAGAUUCAACGGCAUCGGCGUGAC CCAGAACGUGCUGUACGAGAACCAGAAGCAGA UCGCCAACCAGUUCAACAAGGCCAUCAGCCAG AUCCAGGAGAGCCUGACCACCACCAGCACCGC CCUGGGCAAGCUGCAGGACGUGGUGAACCAGA ACGCCCAGGCCUGAACACCCUGGUGAAGCAG CUGAGCAGCAACUUCGGCGCCAUCAGCAGCGU GCUGAACGACAUCCUGAGCAGACUGGACCCUC CUGAGGCCGAGGUGCAGAUCGACAGACUGAUC ACCGGCAGACUGCAGAGCCUGCAGACCUACGU GACCCAGCAGCUGAUCAGAGCCGCCGAGAUCA GAGCCAGCGCCAACCUGGCCGCCACCAAGAUG AGCGAGUGCGUGCUGGGCCAGAGCAAGAGAG UGGACUUCUGCGGCAAGGGCUACCACCUGAUG AGCUUCCCUCAGGCCGCCCCUCACGGCGUGG UGUUCCUGCACGUGACCUACGUGCCUAGCCAG GAGAGAAACUUCACCACCGCCCCUGCCAUCUG CCACGAGGGCAAGGCCUACUUCCCUAGAGAGG GCGUGUUCGUGUUCAACGGCACCAGCUGGUU CAUCACCCAGAGAAACUUCUUCAGCCCUCAGA UCAUCACCACCGACAACACCUUCGUGAGCGGC AGCUGCGACGUGGUGAUCGGCAUCAUCAACAA CACCGUGUACGACCCUCUGCAGCCUGAGCUGG ACAGCUUCAAGGAGGAGCUGGACAAGUACUUC AAGAACCACACCAGCCCUGACGUGGACCUGGG CGACAUCAGCGGCAUCAACGCCAGCGUGGUGA ACAUCCAGAAGGAGAUCGACAGACUGAACGAG	
--	--	--

	GUGGCCAAGAACCUGAACGAGAGCCUGAUCGA CCUGCAGGAGCUGGGCAAGUACGAGCAGUACA UCAAGUGGCCUUGGUACGUGUGGCUGGGCUU CAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGA CCAUCCUGCUGUGCUGCAUGACCAGCUGCUGC AGCUGCCUGAAGGGCGCCUGCAGCUGCGGCA GCUGCUGCAAGUUCGACGAGGACGACAGCGAG CCUGUGCUGAAGGGCGUGAAGCUGCACUACAC C	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFIFLFLTLTSGSDLESCCTTFDDVQAPNYPQHSS SRRGVYYPDEIFRSDTLYLTDLFLPFYSNVTGFH TINHRFDNPVIPFKDGVYFAATEKSNVVRGWVFGS TMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVS KPTGTQTHTMIFDNAFNCTFEYISDSFSLDVAEKS GNFKHLREFVFKNKDGLYVYKGYQPIDVVRDLPS GFNILKPIFKLPLGINITNFRAILTAFLPAQDTWGTS AAYFVGYLKPATFMLKYDENGITDAVDCSQNPLA ELKCSVKSFEIDKGIYQTSNFRVAPSKEVVRFPNIT NLCPFGEVFNATTFPSVYAWERKRISNCVADYSVL YNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVVK GDDVRQIAPGQTGVIADYNYKLPDDFTGCVLAWN TRNIDATQTGNVNYKYRSLRHGKLRPFERDISNVP FSPDGKPCTPPAFNCYWPLNDYGFYITNGIGYQPY RVVLSFELLNAPATVCGPKLSTD LIKNQCVNFNF NGLTGTGVLTPSSKRFQPFQQFGRDVLDFDTSVR DPKTSEILDISPCSFGGVSVITPGTNTSSEVAVLYQ DVNCTDVPVAIHADQLTPSWRVYSTGNNVFQTQA GCLIGAEHVDTSYECDIPIGAGICASYHTVSSLRST	49

	SQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTE VMPVSMAKTSVDCNMYICGDSTECANLLLQYGSF CTQLNRALSGIAVEQDRNTREVFAQVKQMYKTPT LKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLA DAGFMKQYGECLGDINARDLICAQKFNGLTVLPPL LTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFA MQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQI QESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSN FGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQT YVTQQLIRAAEIRASANLAATKMSECVLGQSKRVD FCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNF TTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFF SPQIITTDNTFVSGSCDVVIGIINNTVYDPLQPELDS FKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEI DRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWL GFIAGLIAIVMTILLCCMTSCCSCCLKGACSCGSCC KFDEDDSEPVLKGVKLHYT	
Cola de PoliA	100 nt	
Variante 1 de la proteína MERS S		
La SEQ ID NO: 60 consiste en el extremo 5' al extremo 3': 5' UTR SEQ ID NO: 2, ARNm ORF SEQ ID NO: 61 y 3' UTR SEQ ID NO: 4.		60
Estructura química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
5' UTR	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de	AUGAUCCACAGCGUGUCCUGCUGAUGUCCU CCUUACCCCUACCGAGAGCUACGUGGACGUCG GCCUGACAGCGUUAAGAGCGCUUGCAUCGAG GUGGACAUCCAGCAGACCUUCUUCGACAAGAC CUGGCCUAGACCUAUCGACGUGAGCAAGGCCG	61

detención)	ACGGCAUCAUCUACCCUCAGGGCAGAACCUAC AGCAACAUCACCAUACCUACCAGGGCCUGUU CCCUUAUCAGGGCGACCACGGCGACAUGUACG UGUACAGCGCCGGCCACGCCACCGGCACAACG CCCCAGAAGCUUUUCGUGGCCAACUACUCCCA GGACGUGAAGCAGUUCGCCAACGGCUUCGUG GUGAGAAUCGGCGCCGCCGCCAACUCCACUGG AACCGUGAUCAUCAGCCCUAGCACCAGCGCCA CCAUCAGAAAGAUUUUAUCCUGCCUUUAUGCUG GGCUCCAGCGUGGGUAACUUCAGCGACGGCAA GAUGGGCAGAUUCUUAACACACCCUGGUGC UGCUGCCUGACGGCUGCGGCACCCUGCUGAG AGCCUUCUACUGCAUCCUGGAGCCUAGAAGCG GCAACCACUGCCCUGCCGGCAACAGCUACACC AGCUUCGCAACCUAUCACACCCCUGCCACCGA CUGUUCUGACGGUAACUACAACAGAAACGCCA GCCUGAACAGCUUCAAGGAGUACUUCAACCUG AGAAACUGCACCUUCAUGUACACCUAUAAUAUC ACCGAGGACGAGAUCUCCUGAGUGGUUCGGCAU AACCCAGACCGCCCAAGGCGUGCACCUGUUCA GCAGCAGAUACGUUGAUCUGUACGGCGGCAAC AUGUUCCAGUUCGCUACCCUGCCUGUGUACGA CACCAUCAAGUACUACAGCAUCAUCCCUCAUUC UAUUAGAAGCAUCCAGAGCGACAGAAAGGCCU GGGCCGCUUUCUACGUUAUACAAGCUGCAGCCU CUCACAUUCUUGCUCGACUUCUCUGUGGACGG CUAUAUCCGCAGGGCCAUCGACUGCGGCUUCA ACGACCUGAGCCAGCUGCACUGCAGCUACGAG AGCUUCGACGUGGAGAGCGGAGUUUAUUCCGU GAGCAGCUUCGAGGCCAAGCCUAGCGGCUCUG UAGUGGAGCAGGCCGAGGGCGUGGAGUGCGA UUUCAGCCCUCUGCUGAGCGGUACCCCUCCUC	
------------	---	--

	AGGUGUACAACUUCAAGAGACUGGUGUUCACG AACUGCAACUACAAUCUGACCAAACUGCUUUC GCUUUUCUCCGUGAACGACUUCACCUGCAGCC AGAUUUCUCCGGCAGCCAUCGCCAGCAACUGC UACAGCAGCUUGAUCCUUGACUACUUCAGCUA CCCUCUGAGCAUGAAGUCCGACUUAAGUGUAU CCUCAGCCGGCCCUAUCAGCCAGUUCAAUUAC AAGCAGAGCUUCAGCAACCCUACCUGCCUAAU UUUGGCCACCGUGCCUCACAACCUGACUACAA UUACCAAGCCACUCAAGUAUUCUACAUAACA AGUGUAGCCGAUUCUGAGCGACGACAGAACC GAGGUGCCUCAGCUGGUGAACGCCAACCAGUA CAGCCCUUGCGUGUCGAUCGUGCCAAGUACCG UGUGGGAGGACGGCGACUACUACAGAAAGCAG CUGUCUCCUCUCGAAGGCGGGGUGGCUGG UGGCAAGCGGAAGCACAGUGGCCAUGACCGAG CAGCUGCAGAUGGGCUUCGGAAUUACCGUGCA GUACGGCACCGACACCAAUAGUGUCUGCCCUA AGCUGGAUUUCGCGAACGACACUAAGAUUGCC UCCCAACUGGGAAAUUGCGUAGAGUACUCUCU GUACGGAGUGUCCGGCAGAGGUGUCUUCAG AAUUGCACAGCCGUGGGCGUGAGACAGCAGAG AUUCGUCUACGACGCCUACCAGAACCUGGUGG GCUAUUAUAGUGACGACGGCAACUACUACUGC CUGCGGGCCUGCGUUAGUGUGCCUGUCUCCG UUAUCUACGACAAGGAGACAAAGACUCACGCC ACACUUUUCGGAUCUGUCGCCUGCGAGCACAU CAGUAGUACCAUGUCUCAGUAUAGCAGAAGCA CCAGGUCUAUGCUGAAGAGACGGGACUCAACC UACGGACCACUUCAGACCCCUUGUGGGCUGCGU GCUGGGCCUCGUAAAUAGCUCUCUGUUUGUG GAGGACUGUAAACUGCCACUGGGCCAGAGCCU	
--	--	--

	GUGUGCUUUACCUGACACACCUAGUACACUGA CACCAGCGAGCGUGGGUAGUGUACCAGGCGA GAUGAGACUGGCCAGCAUCGCUUUCAAUACC CUAUCCAGGUGGACCAGCUCAAUUCCUCUUA UUCAAGCUGAGCAUCCCUACCAAUUUCUCUUU CGGCGUGACCCAGGAGUACAUCCAGACCACAA UACAGAAGGUGACCGUAGAUUGCAAGCAGUAC GUGUGUAACGGAUUCCAGAAGUGCGAGCAAUU GCUCAGGGAGUACGGCCAGUUCUGUAGCAAGA UCAACCAGGCUCUGCACGGGGCCAAUCUGCGA CAGGACGACAGCGUAAGAAACCUGUUCGCCAG CGUAAAGUCUAGCCAGUCGAGUCCAAUCAUAC CAGGCUUCGGCGGAGAUUUCAAUCUCACCUUA UUGGAGCCAGUUUCCAUCUCUACGGGUUCGAG GAGCGCUAGGAGCGCAAUCGAGGACCUGCUGU UCGAUAAGGUCACCAUCGCCGACCCUGGCUAC AUGCAGGGCUACGACGACUGCAUGCAGCAGGG CCCAGCCUCCGCCAGAGAUCUGAUCUGCGCCC AGUACGUCGCCGGCUACAAGGUGCUGCCUCCU CUGAUGGACGUUAACAUGGAGGCCGCCUAUAC UAGUAGUCUUCUGGGAAGCAUUGCAGGCGUG GGCUGGACCGCCGGCCUGUCUAGCUUCGCGG CCAUACCUUUCGCCCAGAGCAUCUUCUACAGA CUGAACGGUGUGGGCAUCACACAACAGGUACU GUCUGAGAAUCAGAAGCUGAUCGCCAACAAGU UCAUCCAGGCACUUGGCGCCAUGCAGACCGGC UUCACCACCACCAACGAGGCCUUCCACAAGGU CCAGGACGCCGUGAACAACAACGCUCAGGCCU UGAGCAAGUUAGCGAGCGAACUUAGCAACACC UUCGGCGCCAUCAGUGCAAGCAUUGGAGACAU UAUCCAGAGGCUCGACCCUCCUGAGCAGGACG CUCAGAUCGAUCGGUUGAUCAACGGCAGACUG	
--	--	--

	ACCACUCUGAACGCCUUCGUUGCCCAACAACU GGUGCGGUCUGAGAGCGCCGCUUUAUCCGCC CAGCUGGCCAAGGACAAGGUUAACGAGUGCGU GAAGGCACAGUCGAAGCGUUCAGGAUUCUGCG GCCAGGGCACCCACAUCGUGAGCUUCGUCGUG AACGCCCCUAACGGCCUGUACUUCAUGCACGU CGGAUAUUACCCUAGCAACCAUAUUGAAGUGG UGAGCGCGUACGGCCUCUGUGACGCAGCUAAU CCUACAAACUGCAUCGCCCCUGUGAACGGUUA CUUCAUCAAGACCAACAACACCAGAAUCGUGGA CGAGUGGUCAUACACGGGCAGUUCAUUCUACG CCCUGAGCCGAUCACUAGCCUUAACACCAAG UACGUGGCCCCACAAGUGACAUACCAGAACAU UAGCACAAACCUGCCUCCACCGCUGUUAGGUA ACAGCACGGGCAUCGACUUCCAGGACGAAUUA GACGAGUUCUUAAGAACGUGUCCACCAGCAU CCCAAACUUCGGCAGCCUGACCCAGAUCAACA CAACCUUACUCGACCUGACCUACGAGAUGCUG AGCCUCCAGCAGGUUGUCAAGGCCCUCAACGA AUCAUAUAUCGACUUGAAGGAGCUUGGCAAUU ACACUUACUACAACAAGUGGCCUUGGUACAUC UGGCUCGGCUUCAUCGCCGGGCUGGUCGCCC UCGCCCUGUGCGUCUUCUUCAUCCUGUGCUG CACAGGUUGUGGAACCAACUGUAUGGGCAAGC UGAAGUGCAACCGUUGCUGUGAUAGAUACGAG GAGUACGAUCUGGAACCACAUAAAGGUGCACGU GCAC	
3' UTR	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de	MIHSVFLLMFLLTPTESYVDVGPDSVKSACIEVDIQ	59

aminoácidos correspondiente	QTFFDKTWPRPIDVSKADGIIYPQGRTYSNITITYQ GLFPYQGDHGD MYVYSAGHATGTTPQKLFVANYS QDVKQFANGFVVRIGAAANSTGTVIISPSTSATIRKI YPAFMLGSSVGNFSDGKMGRFFNHTLVLLPDGCG TLLRAFYCILEPRSGNHCPAGNSYTSFATYHTPAT DCSDGNYNRNASLNSFKEYFNLRNCTFMYTYNITE DEILEWFGITQTAQGVHLFSSRYVDLYGGNMFQFA TLPVYDTIKYYSIIPHSIRSISQSDRKAWAAFYVYKLO PLTFLLDFSVDGYIRRAIDCGFNDLSQLHCSYESFD VESGVYSVSSFEAKPSGSVVEQAEGVECDFSPLL SGTTPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFT CSQISPAAIASNCYSSLILDYFSYPLSMKSDLSVSS AGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKY SYINKCSRFLSDDRTEVPQLVNANQYSPCVSIVPS TVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTE QLQMGFGITVQYGTDTNSVCPKLEFANDTKIASQL GNCVEYSLYGVSGRGVFQNCTAVGVRQQRFFVYD AYQNLVGYYSDDGNYCLRACVSVPVSVIYDKET KTHATLFGSVACEHISSTMSQYSRSTRSMLKRRD STYGPLQTPVGCVLGLVNSSLFVEDCKLPLGQSLC ALPDTPSTLTPASVGSVPGEMRLASIAFNHPIQVD QLNSSYFKLSIPTNFSFGVTQEYIQTTIQKVTVDCK QYVCNGFQKCEQLLREYGQFC SKINQALHGANLR QDDSVRNLFASVKSSQSSPIIPGFGGDFNLTLLEP VSISTGSRARSARSAIEDLLFDKVTIADPGYMQGYDD CMQQGPASARDLICAQYVAGYKVLPLMDVNMEA AYTSSLLGSIAGVGWTAGLSSFAAIPFAQSIFYRLN GVGITQQVLSENQKLIANKFNQALGAMQTGFTTTN EAFHKVQDAVNNNAQALSKLASELSNTFGAISASI GDIIQRLDPPEQDAQIDRLINGRLTTLNAFVAQQLV RSESAALSAQLAKDKVNECVKAQSKRSGFCGQGT HIVSFVVNAPNGLYFMHVGYYPSNHIEVVSAYGLC	
--------------------------------	---	--

	DAANPTNCIAPVNGYFIKTNNTRIVDEWSYTGSSF YAPEPITSLNTKYVAPQVTYQNISTNLPPPLLGNST GIDFQDELDEFFKNVSTSIPNFGSLTQINTTLLDLTY EMLSLQQVVKALNESYIDLKELGNYTYYNKWPWYI WLGFIAGLVALALCVFFILCCTGCGTNCMGKLCN RCCDRYEEYDLEPHKVHVH	
Cola de PoliA	100 nt	
Variante 25 del SARS-CoV-2		
SEQ ID NO: 86 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 62, y UTR 3' SEQ ID NO: 4.		86
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
UTR 5'	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG	62

	GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCCUUGCAACGGCGUGGAGGGCUUCAAUCGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC	
--	---	--

	CCCAGCCACCGUGUGUGUGGCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCGG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCAUCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCGUGU CACUGAGGAGCGUGGCCAGCCAGAGCAUCAUC GCCUACACCAUGAGCCUGGGCGCCGAGAACAG CGUGGCCUACAGCAACAACAGCAUCGCCAUCC CCACCAACUUCACCAUCAGCGUGACCACCGAG AUUCUGCCCGUGAGCAUGACCAAGACCAGCGU GGACUGCACCAUGUACAUCUGCGGCGACAGCA CCGAGUGCAGCAACCUGCUGCUGCAGUACGGC AGCUUCUGCACCCAGCUGAACCGGGCCCUGAC CGGCAUCGCCGUGGAGCAGGACAAGAACACCC AGGAGGUGUUCGCCAGGUGAAGCAGAUCUAC AAGACCCCUCCAUCAAGGACUUCGGCGGCUU CAACUUCAGCCAGAUCUGCCCGACCCAGCA AGCCCAGCAAGCGGAGCUUCAUCGAGGACCUG CUGUUCAACAAGGUGACCCUAGCCGACGCCGG CUUCAUCAAGCAGUACGGCGACUGCCUCGGCG ACAUAGCCGCCCGGGACCUGAUCUGCGCCAG AAGUUCAACGGCCUGACCGUGCUGCCUCCCU	
--	---	--

	GCUGACCGACGAGAUGAUCGCCAGUACACCA GCGCCCUGUUAGCCGGAACCAUCACCAGCGGC UGGACUUUCGGCGCUGGAGCCGCUCUGCAGA UCCCCUUCGCCAUGCAGAUUGGCCUACCGGUUC AACGGCAUCGGCGUGACCCAGAACGUGCUGUA CGAGAACCAGAAGCUGAUCGCCAACCAGUUCA ACAGCGCCAUCGGCAAGAUCAGGACAGCCUG AGCAGCACCGCUAGCGCCCUGGGCAAGCUGCA GGACGUGGUGAACCAGAACGCCAGGCCUGA ACACCCUGGUGAAGCAGCUGAGCAGCAACUUC GGCGCCAUCAGCAGCGUGCUGAACGACAUCCU GAGCCGGCUGGACAAGGUGGAGGCCGAGGUG CAGAUCGACCGGCUGAUCACUGGCCGGCUGCA GAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCCGCGGAGAUUCGGGCCAGCGCCAA CCUGGCCGCCACCAAGAUGAGCGAGUGCGUGC UGGGCCAGAGCAAGCGGGUGGACUUCUGCGG CAAGGGCUACCACCUGAUGAGCUUUCGCCAGA GCGCACCCACGGAGUGGUGUUCUGCACGU GACCUACGUGCCCGCCCAGGAGAAGAACUUCA CCACCGCCCCAGCCAUCUGCCACGACGGCAAG GCCACUUUCCCCGGGAGGGCGUGUUCGUGA GCAACGGCACCCACUGGUUCGUGACCCAGCGG AACUUCUACGAGCCCCAGAUCAUCACCACCGA CAACACCUUCGUGAGCGGCAACUGCGACGUGG UGAUCGGCAUCGUGAACAACACCGUGUACGAU CCCCUGCAGCCCGAGCUGGACAGCUUCAAGGA GGAGCUGGACAAGUACUUCAAGAAUCACACCA GCCCCGACGUGGACCUGGGCGACAUCAGCGG CAUCAACGCCAGCGUGGUGAACAUCAGAAAGG AGAUCGAUCGGCUGAACGAGGUGGCCAAGAAC CUGAACGAGAGCCUGAUCGACCUGCAGGAGCU	
--	--	--

	GGGCAAGUACGAGCAGUACAUCAAGUGGCCCU GGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUG	
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUCCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSP TKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTL EILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGA EHVNNSYECDIPI GAGICASYQTQTVSLRSVASQSIIAYTMSLGAENS VAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMY ICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKP SKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARD LICAQKFNGLTVLPPLLTD EMI AQYTSALLAGTITSG	63

	WTFGAGAALQIPFAMQPMAYRFNGIGVTQNVLYEN QKLIANQFNNSAIGKIQDSLSTASALGKLQDVVNQN AQALNTLVKQLSSNFGAISSVLNDILSRDKVEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK MSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVV FLHVTYVPAQEKNFTTAPAICHGDKAHFPREGVHV SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGI VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMVTIML	
Cola de PoliA	100 nt	
Variante 26 del SARS-CoV-2		
SEQ ID NO: 87 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 64, y UTR 3' SEQ ID NO: 4.		87
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
UTR 5'	GGGAAAUAAAGAGAGAAAAGAAGAGUAAGAAGAA AUUAUAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC	64

	CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA	
--	--	--

	CCCC UUGCAACGGCGUGGAGGGCUUCAAACUGC UACU UCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGU UCCAGACCCGGGCGG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACA UCCCCAU CGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCGUGU CACUGAGGAGCGUGGCCAGCCAGAGCAUCAUC GCCUACACCAUGAGCCUGGGCGCCGAGAACAG CGUGGCCUACAGCAACAACAGCAUCGCCAUCC CCACCAACUUCACCAUCAGCGUGACCACCGAG AUUCUGCCCGUGAGCAUGACCAAGACCAGCGU GGACUGCACCAUGUACAUCUGCGGCGACAGCA CCGAGUGCAGCAACCUGCUGCUGCAGUACGGC AGCUUCUGCACCCAGCUGAACCGGGCCCUGAC CGGCAUCGCCGUGGAGCAGGACAAGAACACCC AGGAGGUGUUCGCCCAGGUGAAGCAGAUCUAC AAGACCCCUCCAUCAAGGACUUCGGCGGGCUU CAACUUCAGCCAGAUCCUGCCCGACCCAGCA AGCCCAGCAAGCGGAGCUUCAUCGAGGACCUG	
--	--	--

	CUGUUCAACAAGGUGACCCUAGCCGACGCCGG CUUCAUCAAGCAGUACGGCGACUGCCUCGGCG ACAUAGCCGCCCGGGACCUGAUCUGCGCCCAG AAGUUCAACGGCCUGACCGUGCUGCCUCCCCU GCUGACCGACGAGAUGAUCGCCCAGUACACCA GCGCCCUGUUAGCCGGAACCAUCACCAGCGGC UGGACUUUCGGCGCUGGAGCCGCUCUGCAGA UCCCCUUCGCCAUGCAGAUGGCCUACCGGUUC AACGGCAUCGGCGUGACCCAGAACGUGCUGUA CGAGAACCAGAAGCUGAUCGCCAACCAGUUCA ACAGCGCCAUCGGCAAGAUCAGGACAGCCUG AGCAGCACCGCUAGCGCCCUGGGCAAGCUGCA GGACGUGGUGAACCAGAACGCCCAGGCCUGA ACACCCUGGUGAAGCAGCUGAGCAGCAACUUC GGCGCCAUCAGCAGCGUGCUGAACGACAUCU GAGCCGGCUGGACAAGGUGGAGGCCGAGGUG CAGAUCGACCGGCUGAUCACUGGCCGGCUGCA GAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCCCGCCGAGAUUCGGGCCAGCGCCAA CCUGGCCGCCACCAAGAUGAGCGAGUGCGUGC UGGGCCAGAGCAAGCGGGUGGACUUCUGCGG CAAGGGCUACCACCUGAUGAGCUUUCGCCAGA GCGCACCCACGGAGUGGUGUUCUGCACGU GACCUACGUGCCCGCCCAGGAGAAGAACUUCA CCACCGCCCCAGCCAUCUGCCACGACGGCAAG GCCACUUUCCCCGGGAGGGCGUGUUCGUGA GCAACGGCACCCACUGGUUCGUGACCCAGCGG AACUUCUACGAGCCCCAGAUCAUCACCACCGA CAACACCUUCGUGAGCGGCAACUGCGACGUGG UGAUCGGCAUCGUGAACAACACCGUGUACGAU CCCCUGCAGCCCGAGCUGGACAGCUUCAAGGA GGAGCUGGACAAGUACUUCAAGAAUCACACCA	
--	--	--

	GCGCCGACGUGGACCUGGGCGACAUCAGCGG CAUCAACGCCAGCGUGGUGAACAUCCAGAAGG AGAUCGAUCGGCUGAACGAGGUGGCCAAGAAC CUGAACGAGAGCCUGAUCGACCUGCAGGAGCU GGGCAAGUACGAGCAGUACAUCAAGUGGCCCCU GGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUGA AGAAGAAGAAGCGGCCACGGAACUCCUACAAG UGCGGCACCAACACCAUGGAGCGGGAGGAGAG CGAGCAGACCAAGAAGCGGGAGAAGAUCCACA UUCCUGAACGGUCCGACGAAGCCCAGCGGGUG UUCAAGAGCAGCAAGACCAGCAGCUGCGACAA GAGCGACACCUGCUUC	
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLF RKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVVL SFELLHAPATVC	65

	GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAHEVNNSYECDIPI GAGICASYQTQTVSLRSVASQSIIAYTMSLGAENS VAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMY ICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKP SKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARD LICAQKFNGLTVLPPLLTDemiaQYTSALLAGTITSG WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYEN QKLIANQFNsAIGKIQDSLSSTASALGKLQDVVNQN AQALNTLVKQLSSNFGAISSVLNDILSRDKVEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK MSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVV FLHVTYVPAQEKNFTTAPAICHGDKAHFPREGVfV SNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGI VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLKKKKRP RNSYKCGTNTMEREESEQTKKREKIHIPERSDEAQ RVFKSSKTSSCDKSDTCF	
Cola de PoliA	100 nt	
Variante 27 del SARS-CoV-2		
SEQ ID NO: 88 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 66, y UTR 3' SEQ ID NO: 4.		88
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NmpNp	
UTR 5'	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG	66

construcción de ARNm (excluyendo el codón de detención)	UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUACCCGGUUCAGACCCUGCUGGGCCUG CACCGGAGCUACCUGACCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU	
---	---	--

	GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAU CGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAU CGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCUUGCAACGGCGUGGAGGGCUUCAAUCGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AAUUC AACGGCCUUAACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCCG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCCAUCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCGUGU CACUGAGGAGCGUGGCCAGCCAGAGCAUCAUC	
--	---	--

	GCCUACACCAUGAGCCUGGGCGCCGAGAACAG CGUGGCCUACAGCAACAACAGCAUCGCCAUCC CCACCAACUUCACCAUCAGCGUGACCACCGAG AUUCUGCCCGUGAGCAUGACCAAGACCAGCGU GGACUGCACCAUGUACAUCUGCGGCGACAGCA CCGAGUGCAGCAACCUGCUGCUGCAGUACGGC AGCUUCUGCACCCAGCUGAACCGGGCCCUGAC CGGCAUCGCCGUGGAGCAGGACAAGAACACCC AGGAGGUGUUCGCCCAGGUGAAGCAGAUCUAC AAGACCCCUCCAUCAAGGACUUCGGCGGCUU CAACUUCAGCCAGAUCUGCCCGACCCAGCA AGCCCAGCAAGCGGAGCUUCAUCGAGGACCUG CUGUUCAACAAGGUGACCCUAGCCGACGCCGG CUUCAUCAAGCAGUACGGCGACUGCCUCGGCG ACAUAGCCGCCCGGGACCUGAUCUGCGCCCAG AAGUUCAACGGCCUGACCGUGCUGCCUCCCCU GCUGACCGACGAGAUGAUCGCCCAGUACACCA GCGCCCUGUUAGCCGGAACCAUCACCAGCGGC UGGACUUUCGGCGCUGGAGCCGCUCUGCAGA UCCCCUUCGCCAUGCAGAUUGGCCUACCGGUUC AACGGCAUCGGCGUGACCCAGAACGUGCUGUA CGAGAACCAGAAGCUGAUCGCCAACCAGUUCA ACAGCGCCAUCGGCAAGAUCAGGACAGCCUG AGCAGCACCGCUAGCGCCCUGGGCAAGCUGCA GGACGUGGUGAACCAGAACGCCAGGCCUGA ACACCCUGGUGAAGCAGCUGAGCAGCAACUUC GGCGCCAUCAGCAGCGUGCUGAACGACAUCU GAGCCGGCUGGACAAGGUGGAGGCCGAGGUG CAGAUCGACCGGCUGAUCACUGGCCGGCUGCA GAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCGCGCGAGAUUCGGGCCAGCGCCAA CCUGGCCGCCACCAAGAUGAGCGAGUGCGUGC	
--	---	--

	UGGGCCAGAGCAAGCGGGUGGACUUCUGCGG CAAGGGCUACCACCUGAUGAGCUUUCGCCAGA GCGCACCCACGGAGUGGUGUCCUGCACGU GACCUACGUGCCCGCCCAGGAGAAGAACUUCA CCACCGCCCCAGCCAUCUGCCACGACGGCAAG GCCCACUUUCCCCGGGAGGGCGUGUUCGUGA GCAACGGCACCCACUGGUUCGUGACCCAGCGG AACUUCUACGAGCCCCAGAUAUCACCACCGA CAACACCUUCGUGAGCGGCAACUGCGACGUGG UGAUCGGCAUCGUGAACAACACCGUGUACGAU CCCUGCAGCCCGAGCUGGACAGCUUCAAGGA GGAGCUGGACAAGUACUUAAGAAUCACACCA GCCCCGACGUGGACCUGGGCGACAUCAGCGG CAUCAACGCCAGCGUGGUGAACAUCCAGAAGG AGAUCGAUCGGCUGAACGAGGUGGCCAAGAAC CUGAACGAGAGCCUGAUCGACCUGCAGGAGCU GGGCAAGUACGAGCAGUACAUAAGUGGCCCU GGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUGA AGCGGCAGUACAAGGACAUGAUGAGCGAGGGA GGACCACCUGGCGCUGAGCCACAG	
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUUCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFWHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS	67

	YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSP TKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGA EHVNNSYECDIPI GAGICASYQTQTVSLRSVASQSIIAYTMSLGAENS VAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMY ICGDSTEC SNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKP SKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARD LICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSG WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYEN QKLIANQFN SAIGKIQDSLSTASALGKLQDVVNQN AQALNTLVKQLSSNFGAISSVLNDILSRDKVEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK MSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVV FLHVTYVPAQEKNFTTAPAICH DGKAHFPREGV FV SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGI VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLKRQYKD MMSEGGPPGAEPQ	
Cola de PoliA	100 nt	
Variante 28 del SARS-CoV-2		

SEQ ID NO: 89 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 68, y UTR 3' SEQ ID NO: 4.		89
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NImpNp	
UTR 5'	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCAGACCCUGCUGGCCUG CACCGGAGCUACCUGACCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC	68

UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCUCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCCUUGCAACGGCGUGGAGGGCUUCAAUCGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AAUUAACGGCCUUAACGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC	
---	--

	AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCCG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAAC AGCUACGAGUGCGACAUCCCCAUCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCGUGU CACUGAGGAGCGUGGCCAGCCAGAGCAUCAUC GCCUACACCAUGAGCCUGGGCGCCGAGAACAG CGUGGCCUACAGCAACAACAGCAUCGCCAUCC CCACCAACUUCACCAUCAGCGUGACCACCGAG AUUCUGCCCGUGAGCAUGACCAAGACCAGCGU GGACUGCACCAUGUACAUCUGCGGGCGACAGCA CCGAGUGCAGCAACCUGCUGCUGCAGUACGGC AGCUUCUGCACCCAGCUGAACCGGGCCCUGAC CGGCAUCGCCGUGGAGCAGGACAAGAACACCC AGGAGGUGUUCGCCCAGGUGAAGCAGAUCUAC AAGACCCCUCCAUCAAGGACUUCGGCGGCUU CAACUUCAGCCAGAUCUGCCCGACCCAGCA AGCCCAGCAAGCGGAGCUUCAUCGAGGACCUG CUGUUCAACAAGGUGACCCUAGCCGACGCCGG CUUCAUCAAGCAGUACGGCGACUGCCUCGGCG ACAUAGCCGCCCGGGACCUGAUCUGCGCCCAG AAGUUCAACGGCCUGACCGUGCUGCCUCCCCU GCUGACCGACGAGAUGAUCGCCAGUACACCA GCGCCUGUUAGCCGGAACCAUACACAGCGGC UGGACUUUCGGCGCUGGAGCCGCUCUGCAGA UCCCCUUCGCCAUGCAGAUAGGCCUACCGGUUC AACGGCAUCGGCGUGACCCAGAACGUGCUGUA CGAGAACCAGAAGCUGAUCGCCAACCAGUUCA ACAGCGCCAUCGGCAAGAUCAGGACAGCCUG AGCAGCACCGCUAGCGCCCUGGGCAAGCUGCA	
--	--	--

	GGACGUGGUGAACCAGAACGCCAGGCCUGA ACACCCUGGUGAAGCAGCUGAGCAGCAACUUC GGCGCCAUCAGCAGCGUGCUGAACGACAUCU GAGCCGGCUGGACCCCCCGAGGCCGAGGUG CAGAU CGACCGGCUGAUCACUGGCCGGCUGCA GAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCGCGCGAGAUUCGGGCCAGCGCCAA CCUGGCCGCCACCAAGAUGAGCGAGUGCGUGC UGGGCCAGAGCAAGCGGGUGGACUUCUGCGG CAAGGGCUACCACCUGAUGAGCUUUCGCCAGA GCGCACCCACGGAGUGGUGUCCUGCACGU GACCUACGUGCCCGCCCAGGAGAAGAACUUCA CCACCGCCCCAGCCAUCUGCCACGACGGCAAG GCCACUUCUCCCCGGGAGGGCGUGUUCGUGA GCAACGGCACCCACUGGUUCGUGACCCAGCGG AACUUCUACGAGCCCCAGAUCAUCACCACCGA CAACACCUUCGUGAGCGGCAACUGCGACGUGG UGAUCGGCAUCGUGAACAACACCGUGUACGAU CCCCUGCAGCCCGAGCUGGACAGCUUCAAGGA GGAGCUGGACAAGUACUUCAAGAAUCACACCA GCCCCGACGUGGACCUGGGCGACAUCAGCGG CAUCAACGCCAGCGUGGUGAACAUC CAGAAGG AGAUCGAUCGGCUGAACGAGGUGGCCAAGAAC CUGAACGAGAGCCUGAUCGACCUGCAGGAGCU GGGCAAGUACGAGCAGUACAUCAAGUGGCCCU GGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUG	
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG	69

aminoácidos correspondiente	VYYPDKVFRSSVLHSTQDLFLPFFSNVTWTFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLF RKS NLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAHEVNNSYECDIPI GAGICASYQTQTVSLRSVASQSIIAYTMSLGAENS VAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMY ICGDSTEC SNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKP SKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARD LICAQKFNGLTVLPPLLTDemiaQYTSALLAGTITSG WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYEN QKLIANQFN SAIGKIQDSLSTASALGKLQDVVNQN AQALNTLVKQLSSNFGAISSVLNDILSRDPPEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK MSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVV FLHVTYVPAQEKNFTTAPAICHGDKAHFPREGV FV SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGI VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG	
--------------------------------	--	--

	DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMTIML	
Cola de PoliA	100 nt	
Variante 29 del SARS-CoV-2		
SEQ ID NO: 90 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 70, y UTR 3' SEQ ID NO: 4.		90
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NImpNp	
UTR 5'	GGGAAAUAGAGAGAAAAGAAGAGUAAGAAGAA AUUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUAAGAACCUGCGGGAG UUCGUGUUAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU	70

	GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCCAGACCCUGCUGGCCUG CACCGGAGCUACCUGACCCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCCUUGCAACGGCGUGGAGGGCUUACAACUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AAUUCACCGGCCUUAACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC	
--	---	--

AGCAGUUCGGCCGGGACAUCGCCGACACCACC
 GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU
 CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG
 UGAGCGUGAUCACCCCAGGCACCAACACCAGC
 AACCAGGUGGGCCGUGCUGUACCAGGACGUGAA
 CUGCACCGAGGUGCCCGUGGCCAUCCACGCC
 GACCAGCUGACACCCACCUGGCGGGUCUACAG
 CACCGGCAGCAACGUGUUCAGACCCGGGCGG
 GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC
 AGCUACGAGUGCGACAUCCCCAUCCGGCGCCGG
 CAUCUGUGCCAGCUACCAGACCCAGACCGUGU
 CACUGAGGAGCGUGGCCAGCCAGAGCAUCAUC
 GCCUACACCAUGAGCCUGGGCGCCGAGAACAG
 CGUGGCCUACAGCAACAACAGCAUCGCCAUCC
 CCACCAACUUCACCAUCAGCGUGACCACCGAG
 AUUCUGCCCGUGAGCAUGACCAAGACCAGCGU
 GGACUGCACCAUGUACAUCUGCGGGCGACAGCA
 CCGAGUGCAGCAACCUGCUGCUGCAGUACGGC
 AGCUUCUGCACCCAGCUGAACCAGGGCCCUGAC
 CGGCAUCGCCGUGGAGCAGGACAAGAACACCC
 AGGAGGUGUUCGCCCAGGUGAAGCAGAUCUAC
 AAGACCCCUCCAUCAAGGACUUCGGCGGGCUU
 CAACUUCAGCCAGAUCCUGCCCGACCCAGCA
 AGCCCAGCAAGCGGAGCUUCAUCGAGGACCUG
 CUGUUCAACAAGGUGACCCUAGCCGACGCCGG
 CUUCAUCAAGCAGUACGGCGACUGCCUCGGCG
 ACAUAGCCGCCCGGGACCUGAUCUGCGCCCAG
 AAGUUCAACGGCCUGACCGUGCUGCCUCCCCU
 GCUGACCGACGAGAUGAUCGCCCAGUACACCA
 GCGCCCUGUUAGCCGGAACCAUCACCAGCGGC
 UGGACUUUCGGCGCUGGAGCCGCUCUGCAGA
 UCCCCUUCGCCAUGCAGAUGGCCUACCGGUUC

	AACGGCAUCGGCGUGACCCAGAACGUGCUGUA CGAGAACCAGAAGCUGAUCGCCAACCAGUUCA ACAGCGCCAUCGGCAAGA UCCAGGACAGCCUG AGCAGCACCGCUAGCGCCCUGGGCAAGCUGCA GGACGUGGUGAACCAGAACGCCCAGGCCUGA ACACCCUGGUGAAGCAGCUGAGCAGCAACUUC GGCGCCAUCAGCAGCGUGCUGAACGACAUCCU GAGCCGGCUGGACCCCCCGAGGCCGAGGUG CAGAU CGACCGGCUGAUCACUGGCCGGCUGCA GAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCGCGCCGAGAUUCGGGCCAGCGCCAA CCUGGCCGCCACCAAGAUGAGCGAGUGCGUGC UGGGCCAGAGCAAGCGGGUGGACUUCUGCGG CAAGGGCUACCACCUGAUGAGCUU UCCCCAGA GCGCACCCCACGGAGUGGUGU UCCUGCACGU GACCUACGUGCCCGCCCAGGAGAAGAACUUCA CCACCGCCCCAGCCAUCUGCCACGACGGCAAG GCCACU U UCCCCGGGAGGGCGUGUUCGUGA GCAACGGCACCCACUGGUUCGUGACCCAGCGG AACUUCUACGAGCCCCAGAUCAUCACCACCGA CAACACCUUCGUGAGCGGCAACUGCGACGUGG UGAUCGGCAUCGUGAACAACACCGUGUACGAU CCCCUGCAGCCCGAGCUGGACAGCUUCAAGGA GGAGCUGGACAAGUACUUCAAGAAUCACACCA GCCCCGACGUGGACCUGGGCGACAUCAGCGG CAUCAACGCCAGCGUGGUGAACA UCCAGAAGG AGAUCGAUCGGCUGAACGAGGUGGCCAAGAAC CUGAACGAGAGCCUGAUCGACCUGCAGGAGCU GGGCAAGUACGAGCAGUACAUCAAGUGGCCCU GGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUGA AGAAGAAGAAGCGGCCACGGAACUCCUACAAG	
--	--	--

	UGCGGCACCAACACCAUGGAGCGGGAGGAGAG CGAGCAGACCAAGAAGCGGGAGAAGAUCCACA UUCCUGAACGGUCCGACGAAGCCCAGCGGGUG UUCAAGAGCAGCAAGACCAGCAGCUGCGACAA GAGCGACACCUGCUUC	
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUUCUGCACCCGUACCCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLALHRS YLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLF RKS NLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPI GAGICASYQTQTVSLRSVASQSIIAYTMSLGAENS VAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMY ICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKP	71

	SKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARD LICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSG WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYEN QKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQN AQALNTLVKQLSSNFGAISSVLNDILSRDPPEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK MSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVV FLHVTYVPAQEKNFTTAPAICHGDKAHFPREGV FV SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGI VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLKKKKRP RNSYKCGTNTMEREESEQTKKREKIHIPERSDEAQ RVFKSSKTSSCDKSDTCF	
Cola de PoliA	100 nt	
Variante 30 del SARS-CoV-2		
SEQ ID NO: 91 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 72, y UTR 3' SEQ ID NO: 4.		91
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NImpNp	
UTR 5'	GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA	72

	GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC	
--	---	--

	CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCUUGCAACGGCGUGGAGGGCUUCAAACUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCCG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCCAUCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCGUGU CACUGAGGAGCGUGGCCAGCCAGAGCAUCAUC GCCUACACCAUGAGCCUGGGCGCCGAGAACAG CGUGGCCUACAGCAACAACAGCAUCGCCAUCC CCACCAACUUCACCAUCAGCGUGACCACCGAG AUUCUGCCCGUGAGCAUGACCAAGACCAGCGU GGACUGCACCAUGUACAUCUGCGGCGACAGCA CCGAGUGCAGCAACCUGCUGCUGCAGUACGGC AGCUUCUGCACCCAGCUGAACCGGGCCCUGAC CGGCAUCGCCGUGGAGCAGGACAAGAACACCC	
--	---	--

	AGGAGGUGUUCGCCCAGGUGAAGCAGAUCUAC AAGACCCCUCCCAUCAAGGACUUCGGCGGCUU CAACUUCAGCCAGA UCCUGCCCCGACCCAGCA AGCCCAGCAAGCGGAGCUUCAUCGAGGACCUG CUGUUCAACAAGGUGACCCUAGCCGACGCCGG CUUCAUCAAGCAGUACGGCGACUGCCUCGGCG ACAUAGCCGCCCCGGGACCUGAUCUGCGCCCAG AAGUUCAACGGCCUGACCGUGCUGCCUCCCCU GCUGACCGACGAGAUGAUCGCCCAGUACACCA GCGCCCUGUUAGCCGGAACCAUCACCAGCGGC UGGACUUUCGGCGCUGGAGCCGCUCUGCAGA UCCCUUCGCCAUGCAGAUGGCCUACCGGUUC AACGGCAUCGGCGUGACCCAGAACGUGCUGUA CGAGAACCAGAAGCUGAUCGCCAACCAGUUCA ACAGCGCCAUCGGCAAGA UCCAGGACAGCCUG AGCAGCACCGCUAGCGCCCUGGGCAAGCUGCA GGACGUGGUGAACCAGAACGCCCAGGCCUGA ACACCCUGGUGAAGCAGCUGAGCAGCAACUUC GGCGCCAUCAGCAGCGUGCUGAACGACAUCU GAGCCGGCUGGACCCCCCGAGGCCGAGGUG CAGAUCGACCGGCUGAUCACUGGCCGGCUGCA GAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCCGCGGAGAUUCGGGCCAGCGCCAA CCUGGCCGCCACCAAGAUGAGCGAGUGCGUGC UGGGCCAGAGCAAGCGGGUGGACUUCUGCGG CAAGGGCUACCACCUGAUGAGCUU UCCCAGA GCGCACCCACGGAGUGGUGU UCCUGCACGU GACCUACGUGCCCGCCCAGGAGAAGAACUUCA CCACCGCCCCAGCCAUCUGCCACGACGGCAAG GCCACUUU UCCCCGGGAGGGCGUGUUCGUGA GCAACGGCACCCACUGGUUCGUGACCCAGCGG AACUUCUACGAGCCCCAGAUCAUCACCACCGA	
--	--	--

	CAACACCUUCGUGAGCGGCAACUGCGACGUGG UGAUCGGCAUCGUGAACAACACCGUGUACGAU CCCCUGCAGCCCGAGCUGGACAGCUUCAAGGA GGAGCUGGACAAGUACUUCAAGAAUCACACCA GCCCCGACGUGGACCUGGGCGACAUCAGCGG CAUCAACGCCAGCGUGGUGAACAUCCAGAAGG AGAUCGAUCGGCUGAACGAGGUGGCCAAGAAC CUGAACGAGAGCCUGAUCGACCUGCAGGAGCU GGGCAAGUACGAGCAGUACAUCAAGUGGCCCCU GGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUGA AGCGGCAGUACAAGGACAUGAUGAGCGAGGGA GGACCACCUGGCGCUGAGCCACAG	
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLF RKS NLKP FERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVVL SFELLHAPATVC	73

	GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAHEVNNSYECDIPI GAGICASYQTQTVSLRSVASQSIIAYTMSLGAENS VAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMY ICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKP SKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARD LICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSG WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYEN QKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQN AQALNTLVKQLSSNFGAISSVLNDILSRDPPEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK MSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVV FLHVTYVPAQEKNFTTAPAICHGDKAHFPREGVFFV SNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGI VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLKRQYKD MMSEGGPPGAEPQ	
Cola de PoliA	100 nt	
Variante 31 del SARS-CoV-2		
SEQ ID NO: 92 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 74, y UTR 3' SEQ ID NO: 4.		92
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NImpNp	
UTR 5'	GGGAAAUAAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG	74

ARNm (excluyendo el codón de detención)	ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGC GGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUC CAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCC CAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUCAACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU	
--	--	--

	GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCUUGCAACGGCGUGGAGGGCUUCAAACUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AAUUAACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCGG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCCAUCCGGCGCCG CAUCUGUGCCAGCUACCAGACCCAGACCGUGU CACUGAGGAGCGUGGCCAGCCAGAGCAUCAUC GCCUACACCAUGAGCCUGGGCGCCGAGAACAG	
--	--	--

	CGUGGCCUACAGCAACAACAGCAUCGCCAUCC CCACCAACUUCACCAUCAGCGUGACCACCGAG AUUCUGCCCGUGAGCAUGACCAAGACCAGCGU GGACUGCACCAUGUACAUCUGCGGGCAGAGCA CCGAGUGCAGCAACCUGCUGCUGCAGUACGGC AGCUUCUGCACCCAGCUGAACCAGGGCCCUGAC CGGCAUCGCCGUGGAGCAGGACAAGAACACCC AGGAGGUGUUCGCCCAGGUGAAGCAGAUUCUAC AAGACCCCUCCCAUCAAGGACUUCGGCGGGCUU CAACUUCAGCCAGAUCUGCCCGACCCAGCA AGCCCAGCAAGCGGAGCUUCAUCGAGGACCUG CUGUUCAACAAGGUGACCCUAGCCGACGCCGG CUUCAUCAAGCAGUACGGCGACUGCCUCGGCG ACAUAGCCGCCCGGGACCUGAUCUGCGCCCAG AAGUUCAACGGCCUGACCGUGCUGCCUCCCCU GCUGACCGACGAGAUGAUCGCCAGUACACCA GCGCCCUGUUAGCCGGAACCAUCACCAGCGGC UGGACUUUCGGCGCUGGAGCCGCUCUGCAGA UCCCCUUCGCCAUGCAGAUGGCCUACCGGUUC AACGGCAUCGGCGUGACCCAGAACGUGCUGUA CGAGAACCAGAAGCUGAUCGCCAACCAGUUCA ACAGCGCCAUCGGCAAGAUCAGGACAGCCUG AGCAGCACCGCUAGCGCCCUGGGCAAGCUGCA GGACGUGGUGAACCAGAACGCCAGGCCUGA ACACCCUGGUGAAGCAGCUGAGCAGCAACUUC GGCGCCAUCAGCAGCGUGCUGAACGACAUCU GAGCCGGCUGGACAAGGUGGAGGCCGAGGUG CAGAUCGACCGGCUGAUCACUGGCCGGCUGCA GAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCCCGCCGAGAUUCGGGCCAGCGCCAA CCUGGCCGCCACCAAGAUGAGCGAGUGCGUGC UGGGCCAGAGCAAGCGGGUGGACUUCUGCGG	
--	---	--

	CAAGGGCUACCACCUGAUGAGCUUUCUCCAGAG GCGCACCCACGGAGUGGUGUUCUGCACGU GACCUACGUGCCCGCCCAGGAGAAGAACUUCA CCACCGCCCCAGCCAUCUGCCACGACGGCAAG GCCACUUCUCCCGGGAGGGCGUGUUCGUGA GCAACGGCACCCACUGGUUCGUGACCCAGCGG AACUUCUACGAGCCCCAGAUCAUCACCACCGA CAACACCUUCGUGAGCGGCAACUGCGACGUGG UGAUCGGCAUCGUGAACAACACCGUGUACGAU CCCCUGCAGCCCGAGCUGGACAGCUUCAAGGA GGAGCUGGACAAGUACUUAAGAAUCACACCA GCCCCGACGUGGACCUGGGCGACAUCAGCGG CAUCAACGCCAGCGUGGUGAACAUCAGGAAGG AGAUCGAUCGGCUGAACGAGGUGGCAAGAAC CUGAACGAGAGCCUGAUCGACCUGCAGGAGCU GGGCAAGUACGAGCAGUACAUAAGUGGCCCCU GGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUGU GCUGCAUGACCAGCUGCUGCAGCUGCCUGAAG GGCUGUUGCAGCUGCGGCAGCUGCUGCAAGU UCGACGAGGACGACAGCGAGCCCGUGCUGAAG GGCGUG	
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUCCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI	75

	NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSP TKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPI GAGICASYQTQTVSLRSVASQSIIAYTMSLGAENS VAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMY ICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKP SKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARD LICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSG WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYEN QKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQN AQALNTLVKQLSSNFGAISSVLNDILSRDKVEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK MSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVV FLHVTYVPAQEKNFTTAPAICHDGKAHFPREGV FV SNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGI VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSC CSCLKGCCSCGSCCKFDEDDSEPVLKGV	
Cola de PoliA	100 nt	
Variante 32 del SARS-CoV-2		

SEQ ID NO: 93 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 76, y UTR 3' SEQ ID NO: 4.		93
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NImpNp	
UTR 5'	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC	76

	UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCUCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCCUUGCAACGGCGUGGAGGGCUUCAAUCGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AAUUAACGGCCUUAACGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGGCG UGAGCGUGAUCACCCAGGCACCAACACCAGC	
--	---	--

	AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCCG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCCAUCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCGUGU CACUGAGGAGCGUGGCCAGCCAGAGCAUCAUC GCCUACACCAUGAGCCUGGGCGCCGAGAACAG CGUGGCCUACAGCAACAACAGCAUCGCCAUCC CCACCAACUUCACCAUCAGCGUGACCACCGAG AUUCUGCCCGUGAGCAUGACCAAGACCAGCGU GGACUGCACCAUGUACAUCUGCGGGCGACAGCA CCGAGUGCAGCAACCUGCUGCUGCAGUACGGC AGCUUCUGCACCCAGCUGAACCAGGGCCCUGAC CGGCAUCGCCGUGGAGCAGGACAAGAACACCC AGGAGGUGUUCGCCCAGGUGAAGCAGAUUCUAC AAGACCCCUCCCAUCAAGGACUUCGGCGGCUU CAACUUCAGCCAGAUCUGCCCGACCCAGCA AGCCCAGCAAGCGGAGCUUCAUCGAGGACCUG CUGUUCAACAAGGUGACCCUAGCCGACGCCGG CUUCAUCAAGCAGUACGGCGACUGCCUCGGCG ACAUAGCCGCCCGGGACCUGAUCUGCGCCCAG AAGUUCAACGGCCUGACCGUGCUGCCUCCCCU GCUGACCGACGAGAUGAUCGCCAGUACACCA GCGCCUGUUAGCCGGAACCAUACACAGCGGC UGGACUUUCGGCGCUGGAGCCGCUCUGCAGA UCCCCUUCGCCAUGCAGAUAGGCCUACCGGUUC AACGGCAUCGGCGUGACCCAGAACGUGCUGUA CGAGAACCAGAAGCUGAUCGCCAACCAGUUCA ACAGCGCCAUCGGCAAGAUCAGGACAGCCUG AGCAGCACCGCUAGCGCCUGGGCAAGCUGCA	
--	---	--

	GGACGUGGUGAACCAGAACGCCAGGCCUGA ACACCCUGGUGAAGCAGCUGAGCAGCAACUUC GGCGCCAUCAGCAGCGUGCUGAACGACAUCU GAGCCGGCUGGACCCCCCGAGGCCGAGGUG CAGAU CGACCGGCUGAUCACUGGCCGGCUGCA GAGCCUGCAGACCUACGUGACCCAGCAGCUGA UCCGGGCGCGCGAGAUUCGGGCCAGCGCCAA CCUGGCCGCCACCAAGAUGAGCGAGUGCGUGC UGGGCCAGAGCAAGCGGGUGGACUUCUGCGG CAAGGGCUACCACCUGAUGAGCUUUCGCCAGA GCGCACCCACGGAGUGGUGUCCUGCACGU GACCUACGUGCCCGCCCAGGAGAAGAACUUCA CCACCGCCCCAGCCAUCUGCCACGACGGCAAG GCCACUUUCCCCGGGAGGGCGUGUUCGUGA GCAACGGCACCCACUGGUUCGUGACCCAGCGG AACUUCUACGAGCCCCAGAUCAUCACCACCGA CAACACCUUCGUGAGCGGCAACUGCGACGUGG UGAUCGGCAUCGUGAACAACACCGUGUACGAU CCCCUGCAGCCCGAGCUGGACAGCUUCAAGGA GGAGCUGGACAAGUACUUCAAGAAUCACACCA GCCCCGACGUGGACCUGGGCGACAUCAGCGG CAUCAACGCCAGCGUGGUGAACAUC CAGAAGG AGAUCGAUCGGCUGAACGAGGUGGCCAAGAAC CUGAACGAGAGCCUGAUCGACCUGCAGGAGCU GGGCAAGUACGAGCAGUACAUCAAGUGGCCCU GGUACAUCUGGCUGGGCUUCAUCGCCGGCCU GAUCGCCAUCGUGAUGGUGACCAUCAUGCUGU GCUGCAUGACCAGCUGCUGCAGCUGCCUGAAG GGCUGUUGCAGCUGCGGCAGCUGCUGCAAGU UCGACGAGGACGACAGCGAGCCCGUGCUGAAG GGCGUG	
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU	4

	UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUUCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEVIRQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLF RKS NLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAHEVNNSYECDIPI GAGICASYQTQTVSLRSVASQSIIAYTMSLGAENS VAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMY ICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDK NTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKP SKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARD LICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSG WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYEN QKLIANQFN SAIGKIQDSLSTASALGKLQDVVNQN AQALNTLVKQLSSNFGAISSVLNDILSRDPPEAEV QIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK	77

	MSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVV FLHVTYVPAQEKNFTTAPAICHDGKAHFPREGV SNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGI VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG KYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSC CSCLKGCCSCGSCCKFDEDDSEPVKGV	
Cola de PoliA	100 nt	
Variante 33 del SARS-CoV-2		
SEQ ID NO: 94 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 78, y UTR 3' SEQ ID NO: 4.		94
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
UTR 5'	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUUCGUGUUCUGGUGCUGCUGCCCCUGG UGAGCAGCCAGUGCGUGAACCUGACCACCCGG ACCCAGCUGCCACCAGCCUACACCAACAGCUU CACCCGGGGCGUCUACUACCCCGACAAGGUGU UCCGGAGCAGCGUCCUGCACAGCACCCAGGAC CUGUUCUGCCCUUCUUCAGCAACGUGACCUG GUUCCACGCCAUCCACGUGAGCGGCACCAACG GCACCAAGCGGUUCGACAACCCCGUGCUGCCC UUCAACGACGGCGUGUACUUCGCCAGCACCGA GAAGAGCAACAUCAUCCGGGGCUGGAUCUUCG GCACCACCCUGGACAGCAAGACCCAGAGCCUG CUGAUCGUGAAUAACGCCACCAACGUGGUGAU CAAGGUGUGCGAGUUCAGUUCUGCAACGACC CCUUCUGGGCGUGUACUACCACAAGAACAAC AAGAGCUGGAUGGAGAGCGAGUUCGGGUGU ACAGCAGCGCCAACAACUGCACCUUCGAGUAC	78

	GUGAGCCAGCCCUUCCUGAUGGACCUGGAGG GCAAGCAGGGCAACUUCAAGAACCUGCGGGAG UUCGUGUUCAAGAACAUCGACGGCUACUUCAA GAUCUACAGCAAGCACACCCCAAUCAACCUGG UGCGGGAUCUGCCCCAGGGCUUCUCAGCCCU GGAGCCCCUGGUGGACCUGCCCAUCGGCAUCA ACAUCACCCGGUUCCAGACCCUGCUGGCCCUG CACCGGAGCUACCUGACCCCAGGCGACAGCAG CAGCGGGUGGACAGCAGGCGCGGCUGCUUAC UACGUGGGCUACCUGCAGCCCCGGACCUUCCU GCUGAAGUACAACGAGAACGGCACCAUCACCG ACGCCGUGGACUGCGCCCUGGACCCUCUGAG CGAGACCAAGUGCACCCUGAAGAGCUUCACCG UGGAGAAGGGCAUCUACCAGACCAGCAACUUC CGGGUGCAGCCCACCGAGAGCAUCGUGCGGU UCCCCAACAUACCAACCUGUGCCCCUUCGGC GAGGUGUUCAACGCCACCCGGUUCGCCAGCGU GUACGCCUGGAACCGGAAGCGGAUCAGCAACU GCGUGGCCGACUACAGCGUGCUGUACAACAGC GCCAGCUUCAGCACCUUCAAGUGCUACGGCGU GAGCCCCACCAAGCUGAACGACCUGUGCUUCA CCAACGUGUACGCCGACAGCUUCGUGAUCCGU GGCGACGAGGUGCGGCAGAUCGCACCCGGCC AGACAGGCAAGAUCGCCGACUACAACUACAAG CUGCCCGACGACUUCACCGGCUGCGUGAUCGC CUGGAACAGCAACAACCUCGACAGCAAGGUGG GCGGCAACUACAACUACCUGUACCGGCUGUUC CGGAAGAGCAACCUGAAGCCCUUCGAGCGGGA CAUCAGCACCGAGAUCUACCAAGCCGGCUCCA CCCUUGCAACGGCGUGGAGGGCUUCAACUGC UACUUCCCUCUGCAGAGCUACGGCUUCCAGCC CACCAACGGCGUGGGCUACCAGCCCUACCGGG	
--	---	--

	UGGUGGUGCUGAGCUUCGAGCUGCUGCACGC CCCAGCCACCGUGUGUGGGCCCCAAGAAGAGCA CCAACCUGGUGAAGAACAAGUGCGUGAACUUC AACUUCAACGGCCUUACCGGCACCGGCGUGCU GACCGAGAGCAACAAGAAAUCCUGCCCUUUC AGCAGUUCGGCCGGGACAUCGCCGACACCACC GACGCUGUGCGGGAUCCCCAGACCCUGGAGAU CCUGGACAUCACCCCUUGCAGCUUCGGCGGCG UGAGCGUGAUCACCCCAGGCACCAACACCAGC AACCAGGUGGCCGUGCUGUACCAGGACGUGAA CUGCACCGAGGUGCCCGUGGCCAUCCACGCC GACCAGCUGACACCCACCUGGCGGGUCUACAG CACCGGCAGCAACGUGUUCAGACCCGGGCCG GUUGCCUGAUCGGCGCCGAGCACGUGAACAAC AGCUACGAGUGCGACAUCCCCAUCGGCGCCGG CAUCUGUGCCAGCUACCAGACCCAGACCAAUU CACCCCGGAGGGCAAGGAGCGUGGCCAGCCA GAGCAUCAUCGCCUACACCAUGAGCCUGGGCG CCGAGAACAGCGUGGCCUACAGCAACAACAGC AUCGCCAUCCCCACCAACUUCACCAUCAGCGU GACCACCGAGAUUCUGCCCGUGAGCAUGACCA AGACCAGCGUGGACUGCACCAUGUACAUCUGC GGCGACAGCACCGAGUGCAGCAACCUGCUGCU GCAGUACGGCAGCUUCUGCACCCAGCUGAACC GGGCCUGACCGGCAUCGCCGUGGAGCAGGA CAAGAACACCCAGGAGGUGUUCGCCAGGUGA AGCAGAUCAACAAGACCCCUCCAUCAAGGAC UUCGGCGGCUUCAACUUCAGCCAGAUCUGCC CGACCCAGCAAGCCCAGCAAGCGGAGCUUCA UCGAGGACCUGCUGUUCAACAAGGUGACCCUA GCCGACGCCGGCUUCAUCAAGCAGUACGGCGA CUGCCUCGGCGACAUAAGCCGCCCGGGACCUGA	
--	---	--

	UCUGCGCCCAGAAGUUCAACGGCCUGACCGUG CUGCCUCCCCUGCUGACCGACGAGAUGAUCGC CCAGUACACCAGCGCCCUGUUAGCCGGAACCA UCACCAGCGGCUGGACUUUCGGCGCUGGAGC CGCUCUGCAGAUCCCCUUCGCCAUGCAGAUGG CCUACCGGUUCAACGGCAUCGGCGUGACCCAG AACGUGCUGUACGAGAACCAGAAGCUGAUCGC CAACCAGUUCAACAGCGCCAUCGGCAAGAUC AGGACAGCCUGAGCAGCACCGCUAGCGCCCUG GGCAAGCUGCAGGACGUGGUGAACCAGAACGC CCAGGCCCUGAACACCCUGGUGAAGCAGCUGA GCAGCAACUUCGGCGCCAUCAGCAGCGUGCUG AACGACAUCCUGAGCCGGCUGGACCCUCCCGA GGCCGAGGUGCAGAUCGACCGGCUGAUCACU GGCCGGCUGCAGAGCCUGCAGACCUACGUGAC CCAGCAGCUGAUCCGGGCGCCGAGAUUCGG GCCAGCGCCAACCUGGCCGCCACCAAGAUGAG CGAGUGCGUGCUGGGCCAGAGCAAGCGGGUG GACUUCUGCGGCAAGGGCUACCACCUGAUGAG CUUUCCCCAGAGCGCACCCACGGAGUGGUGU UCCUGCACGUGACCUACGUGCCCGCCCAGGAG AAGAACUUCACCACCGCCCCAGCCAUCUGCCA CGACGGCAAGGCCACUUUCCCCGGGAGGGC GUGUUCGUGAGCAACGGCACCCACUGGUUCGU GACCCAGCGGAACUUCUACGAGCCCCAGAUCA UCACCACCGACAACACCUUCGUGAGCGGCAAC UGCGACGUGGUGAUCGGCAUCGUGAACAACAC CGUGUACGAUCCCCUGCAGCCCGAGCUGGACA GCUUCAAGGAGGAGCUGGACAAGUACUUCAAG AAUCACACCAGCCCCGACGUGGACCUGGGCGA CAUCAGCGGCAUCAACGCCAGCGUGGUGAACA UCCAGAAGGAGAUCGAUCGGCUGAACGAGGUG	
--	--	--

	GCCAAGAACCUGAACGAGAGCCUGAUCGACCU GCAGGAGCUGGGCAAGUACGAGCAGUACAUCA AGUGGCCCUGGUACAUCUGGCUGGGCUUCAU CGCCGGCCUGAUCGCCAUCGUGAUGGUGACCA UCAUGCUGUGCUGCAUGACCAGCUGCUGCAGC UGCCUGAAGGGCUGUUGCAGCUGCGGCAGCU GCUGCAAGUUCGACGAGGACGACAGCGAGCCC GUGCUGAAGGGCGUG	
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRG VYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFL GVYYHKNNKSWMESEFRVYSSANNCTFEYVSQP FLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRS YLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNE NGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNITNLCPFGEVFNATRFASVYA WNRKRISNCVADYSVLVNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYN YKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLF RKS NLKPFERDISTEIYQAGSTPCNGVEGFNCYFP LQSYGFQPTNGVGYQPYRVVLSFELLHAPATVC GPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVS VITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTP TWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPI GAGICASYQTQTNSPRRARSVASQSIIAYTMSLGA	79

	ENSVAYSNNNSIAIPTNFTISVTTEILPVSMTKTSVDC TMYICGDSTEC SNLLLQYGSFCTQLNRALTGIAVE QDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD PSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDI AARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAG TITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNV LYENQKLIANQFNSAIGKIQDSLSTASALGKLQDV VNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDPP EAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQSAP HGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPRE GV FVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCD VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCC MTSCCCLKGCCSCGSCCKFDEDDSEPV LKGV	
Cola de PoliA	100 nt	
Proteína de membrana CoV2		
SEQ ID NO: 95 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 80, y UTR 3' SEQ ID NO: 4.		95
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NlmpNp	
UTR 5'	GGGAAUAAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGGCCGACAGCAACGGCACCAUCACCGUGGA GGAGCUGAAGAAGCUGCUGGAGCAGUGGAACC UGGUGAUCGGCUUCCUGUUCUGACCUGGAU CUGCCUGCUGCAGUUCGCCUACGCCAACCGGA ACCGUUUCUGUACAUCAUCAAGCUGAUCUUC CUGUGGCUGCUGUGGCCCGUGACCCUGGCCU GCUUCGUGCUGGCCGCGCGUGUACCGGAUCAA	80

	CUGGAUCACCGGCGGCAUCGCCAUCGCCAUGG CCUGCCUGGUGGGCCUGAUGUGGCUGAGCUA CUUCAUCGCCAGCUUCCGGCUGUUCGCCCCGGA CCCGGAGCAUGUGGAGCUUCAACCCCGAGACC AACAUCCUGCUGAACGUGCCCCUGCACGGCAC AAUCCUGACCCGGCCCCUGCUGGAGAGCGAGC UUGUGAUCGGCGCCGUGAUCCUGCGGGGCCA CCUGAGGAUCGCCGGCCAUCACCUGGGCCGG UGCGACAUCAAGGACCUGCCCAAGGAGAUCAC CGUGGCCACCAGCCGGACCCUGAGCUACUACA AACUGGGCGCCAGCCAGAGAGUGGCCGGAGAC AGCGGCUUCGCCGCCUACAGCCGGUACCGGAU CGGCAACUACAAGCUGAACACCGACCACAGCA GCAGCAGCGACAACAUCGCCCUGCUGGUGCAG	
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUCCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQ FAYANRNRFLYIIKLIFLWLLWPVTLACFVLAAYRI NWITGGIAIAMACLVGLMWLSYFIASFRLFARTRSM WSFNPETNILLNVPLHGTILTRPLLESELVIGAVILR GHLRIAGHHLGRCDIKDLPKEITVATSRTLSTYYKLG ASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSDNI ALLVQ	81
Cola de PoliA	100 nt	
Proteína de envoltura CoV2		
SEQ ID NO: 96 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 82, y UTR 3' SEQ ID NO: 4.		96
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NImpNp	

UTR 5'	GGGAAAUAAAGAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGUACAGCUUCGUGAGCGAGGAGACCGGCAC CCUGAUCGUGAACAGCGUGCUGCUGUUCUG GCCUUCGUGGUGUUCUGCUGGUGACCCUGG CCAUCCUGACCGCCCUGCGGCUGUGUGCCUAC UGCUGCAACAUCGUGAACGUGAGCCUGGUGAA GCCCAGCUUCUACGUGUACAGCCGGGUGAAGA ACCUGAACAGCAGCCGGGUGCCUGACCUGCUG GUG	82
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUUCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MYSFVSEETGTLIVNSVLLFLAFVVFLVTLAILTAL RLCAYCCNIVNVSLVKPSFYVYSRVKNLNSSRVPD LLV	83
Cola de PoliA	100 nt	
Proteína de nucleocápside CoV2		
SEQ ID NO: 97 consiste en el extremo 5' al extremo 3': UTR 5' SEQ ID NO: 2, ORF de ARNm SEQ ID NO: 84, y UTR 3' SEQ ID NO: 4.		97
Química	1-metilpseudouridina	
Cubierta	7mG(5')ppp(5')NImpNp	
UTR 5'	GGGAAAUAAAGAGAGAGAAAAGAAGAGUAAGAAGAA AUAUAAGACCCCGGCGCCGCCACC	2
ORF de la construcción de ARNm (excluyendo el codón de detención)	AUGAGCGACAACGGCCCUAGAAACCAGCGGAA CGCACCCCGGAUCACCUUUGGCGGCCCCAGCG AUAGCACCGGCAGCAACCAGAACGGCGAGAGA UCAGGGGCCCCGGAGCAAGCAGCGGCGUCCUC AGGGCCUGCCCAACAACACCGCCAGCUGGUUC ACCGCCCUGACCCAGCACGGCAAGGAGGACCU	84

	GAAGUUCCCUCGGGGCCAAGGAGUGCCCAUCA ACACCAACAGCAGCCCCGACGACCAGAUCCGGC UACUACAGAAGGGCCACCCGGAGGAUCCGGGG AGGGGACGGCAAGAUGAAGGACCUGUCUCCCC GGUGGUACUUCUACUAUCUUGGCACGGGCCCCU GAAGCUGGCCUGCCGUACGGCGCAAACAAGGA CGGCAUCAUCUGGGUCGCCACCGAGGGAGCC CUGAACACCCCGAAGGACCACAUCGGCACCCG GAAUCCCGCCAACAACGCCGCCAUCGUUCUGC AGCUGCCCCAGGGCACCACCCUGCCCAAGGGC UUCUACGCCGAGGGCAGCAGAGGCGGCUCACA GGCCAGCAGCCGGUCAAGCAGCCGGAGCCGG AACAGCAGCCGGAACUCCACACCCGGCUCUAG CCGAGGCACAAGCCCCGCCAGAAUGGCAGGAA ACGGCGGGCAGCGCUGCCUUAGCCCUGCUGUU GCUGGACCGGCUGAACCAGCUCGAGAGCAAGA UGAGCGGCAAGGGUCAGCAGCAGCAAGGCCAA ACCGUGACCAAGAAGAGCGCCGCCGAGGCUAG CAAGAAGCCCCGGCAGAAGCGGACCGCCACCA AGGCCUACAACGUGACCCAGGCCUUCGGUCCG AGAGGCCCCGAGCAGACCCAGGGCAACUUCGG CGACCAGGAGCUGAUCCGGCAGGGCACCGACU ACAAGCACUGGCCCCAGAU CGCCAGUUCGCC CCUAGCGCCUCAGCCUUCUUCGGCAUGAGCCG GAUCGGCAUGGAGGUGACUCCCAGCGGCACCU GGCUGACCUACACCGGCGCCAUCAAGCUGGAC GACAAGGACCCCAACUUCAAGGACCAGGUGAU CCUGCUGAACAAGCACAU CGACGCCUACAAGA CCUUUCCGCCACCGAGCCCAAGAAGGACAAG AAGAAGAAGGCCGACGAGACCCAGGCCUGCC CCAACGGCAGAAGAAGCAGCAGACCGUCACCU UACUGCCCGCAGCCGACCUGGACGACUUCAGC	
--	---	--

	AAGCAGCUGCAGCAGAGCAUGAGCAGCGCCGA CAGCACCCAGGCC	
UTR 3'	UGAUAAUAGGCUGGAGCCUCGGUGGCCUAGCU UCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCUUCCUGCACCCGUACCCCGUGGUCUU UGAAUAAAGUCUGAGUGGGCGGC	4
Secuencia de aminoácidos correspondiente	MSDNGPQNQRNAPRITFGGPSDSTGSNQNGERS GARSKQRRPQGLPNNTASWFTALTQHGKEDLKFP RGQGVPIINTNSSPDDQIGYYRRATRRIRGGDGKM KDLSPRWYFYLLGTGPEAGLPYGANKDGIWVATE GALNTPKDHIGTRNPANNAIVLQLPQGTTLPKGF YAEGRGGSQASSRSSSRNSSRNSTPGSSRG TSPARMAGNGGDAALALLLDRLNQLESKMSGKG QQQQGQTVTKKSAAEASKKPRQKRTATKAYNVT QAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQI AQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAIKL DDKDPNFKDQVILLNKHIDAYKTFPTEPKKDKKKK ADETQALPQRQKKQQTVTLLPAADLDDFSKQLQQ SMSSADSTQA	85
Cola de PoliA	100 nt	

* Debe entenderse que cualquiera de los marcos de lectura abiertos y/o secuencias de aminoácidos correspondientes descritas en la Tabla 1 pueden incluir o excluir la secuencia señal. También debe entenderse que la secuencia de señal puede ser reemplazada por una secuencia de señal diferente, por ejemplo, cualquiera de las SEQ ID NO: 38-43.

EQUIVALENTES

Todas las referencias, patentes y solicitudes de patente descritas en este documento se incorporan como referencia con respecto al tema por el que se cita cada una, que en algunos casos puede abarcar la totalidad del documento.

- Se debería entender que los artículos indefinidos "un" y "una", tal como se utilizan en la presente memoria descriptiva y en las reivindicaciones, significan "al menos uno/a", a menos que se indique claramente lo contrario. También debe entenderse que, a menos que se

indique claramente lo contrario, en cualquier método reivindicado en este documento que incluya más de un paso o acción, el orden de los pasos o acciones del método no se limita necesariamente al orden en el que los pasos o las acciones del método se encuentran especificados.

- 5 En las reivindicaciones, así como en la memoria descriptiva anterior, todas las frases transitorias como "que comprende", "que incluye", "que conlleva", "que tiene", "que contiene", "que implica", "que posee", "que se compone de" y similares deben entenderse como abiertas, es decir, que incluyen pero de modo no taxativo. Solo las frases transicionales "que consiste en" y "que consiste esencialmente en" serán frases
- 10 transicionales cerradas o semicerradas, respectivamente, tal como se establece en el Manual del Procedimiento de Examen de Patentes de la Oficina de Patentes de los Estados Unidos, sección 2111.03.

Los términos "aproximadamente" y "sustancialmente" que preceden a un valor numérico significan $\pm 10\%$ del valor numérico indicado.

- 15 Cuando se proporciona un rango de valores, cada valor entre los extremos superior e inferior del rango se contempla y describe específicamente en el presente documento.

Todo el contenido de las solicitudes internacionales números PCT/US2015/02740, PCT/US2016/043348, PCT/US2016/043332, PCT/US2016/058327, PCT/US2016/058324, PCT/US2016/058314, PCT/US2016/058310, PCT/US2016/058321, PCT/US2016/058297,

20 PCT/US2016/058319, y PCT/US2016/058314 se incorporan a la presente mediante referencia.