



UNIVERSIDADE DA CORUÑA

Facultade de Ciencias

## **Grao en Bioloxía**

### **Memoria do Traballo de Fin de Grao**

**Coronavirus: una revisión bibliográfica de su ciclo viral y expresión**

**Coronavirus: unha revisión bibliográfica do seu ciclo viral e expresión**

**Coronavirus: a literature review of their viral cycle and expression**

**Lucía Lema Mira**

**Curso: 2020 - 2021. Convocatoria: Junio**

*Directora académica: María Ángeles Freire Picos*

# Coronavirus: una revisión bibliográfica de su ciclo viral y expresión

## Resumen

A lo largo de esta revisión se estudian los virus de la familia *Coronaviridae*, prestando especial atención a su expresión génica y ciclo viral. Estos virus están ganando importancia médica en las últimas décadas debido a brotes causados por alguno de ellos, como es el caso de la pandemia que se inició en 2019 por el SARS-CoV-2. Para conocerlos, estudiaremos las proteínas que codifica su genoma y las formas que tienen de expresarlas, mediante mecanismos como el *ribosomal frameshifting* además de la obtención de RNAs subgenómicos generados por transcripción discontinua. A su vez, comentaremos estrategias para su estudio como la combinación de las técnicas de RiboSeq y RNASeq, en el análisis de la eficacia de la traducción. Estas investigaciones han adquirido enorme relevancia internacional para la búsqueda de tratamientos, buscándose en estos momentos puntos débiles de estos virus para hacerles frente.

## Resumo

Ao longo desta revisión estúdanse os virus da familia *Coronaviridae*, prestando especial atención a súa expresión xénica e ciclo viral. Estes virus están a ganar importancia médica nas últimas décadas debido a brotes causados por algúns deles, como é o caso da pandemia que comezou en 2019 polo SARS-CoV-2. Para coñecelos, estudaremos as proteínas que codifica o seu xenoma e as maneiras que teñen de expresalas, mediante mecanismos como o *ribosomal frameshifting* ademais da obtención de RNAs subxenómicos xerados por transcripción descontinua. Ao mesmo tempo, falaremos de estratexias para o seu estudo coma a combinación das técnicas de RiboSeq e RNASeq nos estudos da eficacia da tradución. Estas investigación adquiriron enorme relevancia internacional para a busca de tratamentos, buscándose nestes momentos puntos débiles destes virus para facerlles fronte.

## Abstract

Throughout this review, viruses from the *Coronaviridae* family are studied, paying special attention to their genomic expression and viral cycle. These viruses are gaining medical importance in the last few decades due to outbreaks caused by some of them, as in the case of SARS-CoV-2 pandemic, initiated in 2019. To get to know them, we will study the proteins codified by their genomes and the ways this occurs, using mechanisms such as ribosomal frameshifting, in addition to subgenomic RNA obtention by discontinuous transcription.

Moreover, we will focus on strategies used to study them, as the combination between RiboSeq and RNASeq in translation efficiency. These investigations have acquired huge global relevance owing to the search for treatments, as right now they are looking for weak points of these viruses to fight them.

**Palabras clave:** coronavirus, ribosomal frameshifting, expresión, transcripción, ribosome profiling, ciclo viral

### **Abreviaturas:**

CDS- *coding sequence* – secuencia codificante

CHX – Cicloheximida

CS – *core sequence* – secuencia central de la región reguladora TRS

CS-B – CS del cuerpo

CS-L – CS de la secuencia líder

ERGIC - compartimento intermedio retículo endoplasmático-Golgi

gRNA – RNA genómico

HAR – Harringtonina

LTM – lactidomicina

NGS – *Next Generation Sequencing* – secuenciado de última generación

nsp – *non structural protein* – proteína no estructural

ORF – *open reading frame* – pauta abierta de lectura

RBD – *receptor binding domain* – dominio de unión al receptor

RE – retículo endoplasmático

RiboSeq – *ribosome profiling* – perfil ribosómico

RNASeq – secuenciación RNA

RTC- *replication transcription complex* – complejo de replicación transcripcional

sgmRNA – mRNA subgenómico

sgRNA – RNA subgenómico

TMPRSS – *transmembrane serine protease* – serin-proteasa transmembrana

TRS –secuencias reguladoras de la transcripción

TRS-L – TRS de la secuencia líder

TRS-B – TRS del cuerpo

UTR – *untranslated región* – región no traducida

# Índice

Resumen/Resumo/Abstract

Abreviaturas

1. Introducción.....	2
2. Objetivos.....	2
3. Metodología.....	3
4. Desarrollo del tema	
4.1 <b>Estructura de los Coronavirus</b> .....	3
4.2 <b>Ciclo viral</b> .....	4
4.3 <b>Estructura genómica y expresión génica en Coronavirus</b> .....	7
4.4 <b>Traducción del genoma viral y cambio en el marco de lectura ribosómico     (<i>ribosomal frameshifting</i>)</b> .....	8
4.5 <b>Proteínas de coronavirus</b> .....	9
4.6 <b>Replicación y transcripción en coronavirus</b> .....	10
4.7 <b>Técnicas para el estudio de coronavirus</b>	
4.7.1 <b>Perfil ribosómico o <i>ribosome profiling</i></b> .....	13
4.7.2 <b>Secuenciado RNA y <i>ribosome profiling</i> en Coronavirus y ejemplos</b> .....	14
5. Conclusiones.....	16
6. Bibliografía.....	17
7. Anexo I.....	23

## 1. Introducción

En la última década, a causa de las enfermedades que acarrearón a la población mundial, los coronavirus han pasado de ser conocidos básicamente por expertos en virología a formar parte del interés general para la población, especialmente en lo relativo a la transmisión de la enfermedad, los posibles tratamientos y las vacunas.

Los virus de la familia *Coronaviridae* pertenecen a la orden *Nidovirales*, dividiéndose en dos subfamilias: *Torovirinae* y *Coronavirinae* (Irigoyen et al., 2016). En la subfamilia *Coronavirinae* encontraremos los géneros *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* y *Deltacoronavirus* (revisado por Nakagawa et al., 2020). En humanos se relacionan generalmente con el resfriado común, pero en las últimas décadas ha habido tres brotes importantes provocados por tres virus del género *Betacoronavirus*, llamados SARS-CoV, MERS-CoV y SARS-CoV-2 cuyas consecuencias han sido más graves (Paules et al., 2020). Este último, el SARS-CoV-2 (SARS por Severe Acute Respiratory Syndrome) provoca la enfermedad Covid-19, una enfermedad ahora pandémica originada en China en el año 2019.

A nivel molecular, los coronavirus, al igual que resto de los virus del orden *Nidovirales*, expresan sus genes a través de múltiples RNAs subgenómicos, que se traducen en diferentes proteínas por el cambio de marco de lectura ribosómico (*ribosome frameshifting*) y siguen un mecanismo de ensamblaje de viriones único entre virus RNA con envuelta (Masters, 2006).

En este trabajo nos proponemos hacer un estudio molecular de los coronavirus sin centrarnos específicamente en el SARS-CoV2, aunque la mayoría de lo tratado se puede aplicar al mismo.

## 2. Objetivos:

El objetivo principal de este trabajo es la realización de una revisión bibliográfica sobre los coronavirus desde el punto de vista molecular. Para esto estudiaremos su ciclo viral y la expresión de su genoma.

### 3. Metodología:

Para realizar esta revisión bibliográfica se llevó a cabo una búsqueda en Google Scholar (<https://scholar.google.com/>) de artículos de revisión e investigación científica sobre el tema, así como en las bases de datos de PubMed (del NIH, <https://pubmed.ncbi.nlm.nih.gov/>). Se seleccionaron aquellos relacionados con la expresión génica y ciclo viral de coronavirus, así como las investigaciones sobre sus proteínas y metodologías específicas para el estudio de virus de esta familia. Se empleó el gestor bibliográfico Mendeley para citar los artículos y componer la bibliografía.

## 4. Desarrollo del tema

### 4.1 Estructura de los Coronavirus

Las partículas de coronavirus son de forma esférica, con un diámetro medio de aproximadamente 100 nm (Fehr & Perlman, 2015; Nakagawa et al., 2020) y con una envuelta que alojará su genoma (RNA). En el exterior presenta unas espículas superficiales, llamadas peplómeros, que se proyectan 17-20 nm desde la superficie del virión (Lai, 1990; Siddell et al., 1982), contando con una base de 10 nm de diámetro (Sugiyama & Amano, 1981). Estas pueden estar acompañadas de otras más pequeñas, de 5-10 nm de largo, que han sido identificadas como la proteína hemaglutinin-esterasa (Hulswit et al., 2016; Masters, 2006).

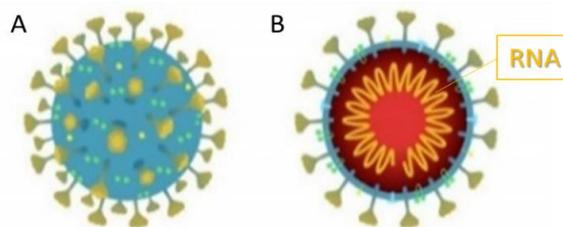


Figura 1 – Estructura general de un coronavirus. A – envuelta . B – interior de la envuelta. Imagen tomada de Freepik (freepik.es).

En su envuelta encontraremos, principalmente, tres proteínas estructurales: las proteínas S, M y E. La proteína M (membrana) es una glicoproteína que sirve para incorporar los componentes virales en nuevos viriones durante la morfogénesis (Neuman et al., 2011). Es la más abundante de las tres y da forma a la envuelta del virión. Mientras, la proteína E (envuelta) es un pequeño polipéptido que se encuentra en bajas cantidades, formando

canales iónicos en la membrana viral y participando en el ensamblaje (Siu et al., 2008; Verdiá-Báguena et al., 2012). La proteína S (espícula) es una glicoproteína que forma las espículas principales e interacciona con la célula huésped iniciando la infección, puesto que se une a receptores de la superficie celular siendo necesaria para la fusión de la membrana viral con la celular (Hulswit et al., 2016).

## 4.2 Ciclo viral

La entrada en las células se inicia con la unión al receptor de la membrana celular y procesamiento proteolítico de la proteína S, que a su vez lleva a la fusión de las membranas (virus-célula). Estos receptores de membrana y sitios de anclaje S suelen ser especie-específicos, habiendo diferentes dominios de unión al receptor (*Receptor Binding Domains*, RBDs) para diferentes virus de esta familia. En el caso de SARS-CoV y SARS-CoV-2, el receptor es ACE2 (la enzima convertidora de angiotensina 2) (Gralinski & Baric, 2015), una peptidasa de la superficie celular que se expresa en la mayoría de los órganos, especialmente en el epitelio del intestino delgado y pulmones (Paules et al., 2020). Después de la unión al receptor ACE2 vírico, las proteínas S son modificadas por la proteasa TMPRSS2/Furina, que causa escisiones y activación en sitios específicos, lo que lleva a la fusión con la membrana (Hulswit et al., 2016) (Figura 2).

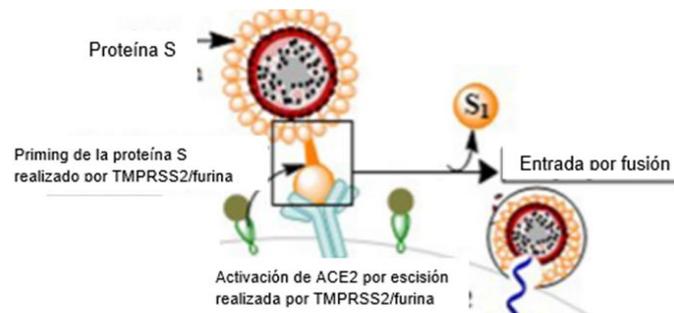


Figura 2 – Interacción de la proteína S con el receptor celular ACE2 (modificada de Senapati et al., 2021) - S1 fragmento escindido de la proteína S.

Así depositan su genoma RNA, cubierto por la proteína N (nucleocápside), en el citoplasma por endocitosis o, directamente, por fusión de membrana en un proceso poco conocido (Hartenian et al., 2020). La proteína N, una fosfoproteína, se asocia al genoma viral en una estructura de collar de cuentas (Fehr & Perlman, 2015) y a su vez se asocia con la proteína M, dirigiendo el empaquetamiento del genoma en nuevas partículas víricas y estabilizando

la nucleocápside (Schoeman & Fielding, 2019). Se considera que protege al RNA viral de ribonucleasas y que, gracias a las interacciones entre proteínas N, se obtiene estabilidad en la nucleocápside. Se ha sugerido un posible rol en la parada de la mitosis, alterando las funciones nucleares de la célula huésped y consiguiendo así un ambiente más favorable para la infección viral (Masters, 2006).

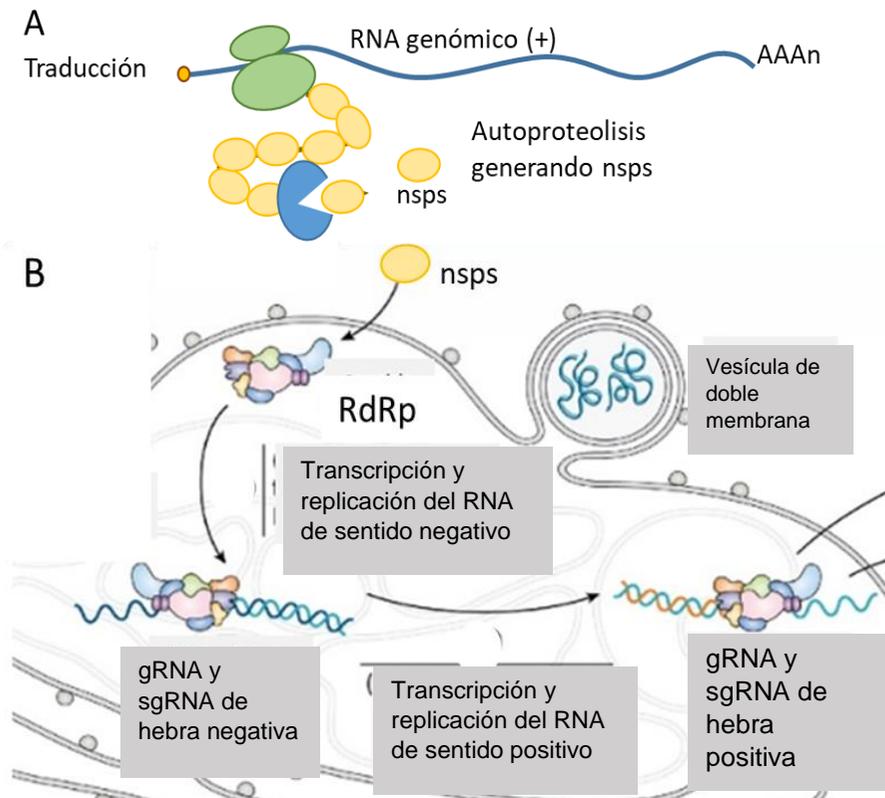


Figura 3 – A. Traducción B. Transcripción y replicación (modificado de Hartenian et al., 2020) – nsps (proteínas no estructurales).

La traducción del RNA genómico y su replicación constituyen la segunda etapa. Este se traducirá a poliproteínas que serán escindidas co-traduccionalmente por proteasas virales (figura 3A) (Hartenian et al., 2020). Las nsps (*non structural proteins*) formarán los componentes del complejo RdRp (RNA polimerasa dependiente de RNA), que usa el RNA genómico (gRNA) como molde para generar RNA genómico y subgenómico de hebra (-). El gRNA (-) sirve como molde para la replicación del genoma. Los RNAs negativos subgenómicos se emplean como molde para la síntesis de RNAs subgenómicos (+) o sgmRNAs por transcripción discontinua (figura 3 B) (aspecto que se estudiará más adelante).

Estos sgmRNAs serán traducidos a proteínas estructurales y accesorias que se asociarán a la membrana del retículo endoplasmático (RE) (Figura 4).

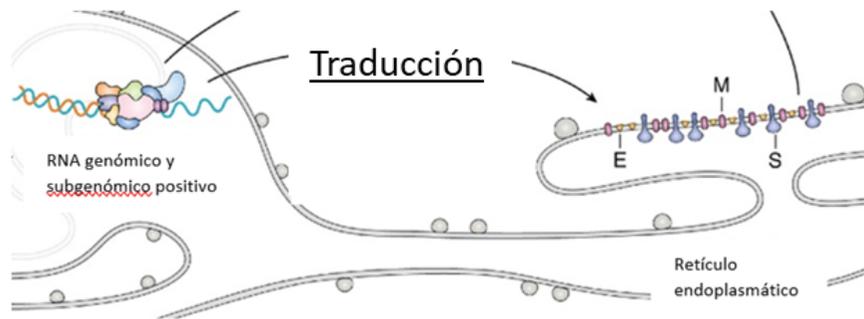


Figura 4 – Traducción de los mRNAs subgenómicos a proteínas estructurales y accesorias. (modificado de Hartenian et al., 2020)

Por último, el gRNA positivo se unirá a la proteína N y entrará al compartimento intermedio entre el retículo endoplasmático y el aparato de Golgi, que estará asociado a las proteínas S, E y M. Así se obtendrá el virión con envuelta, que saldrá de la célula por exocitosis (Figura 5).

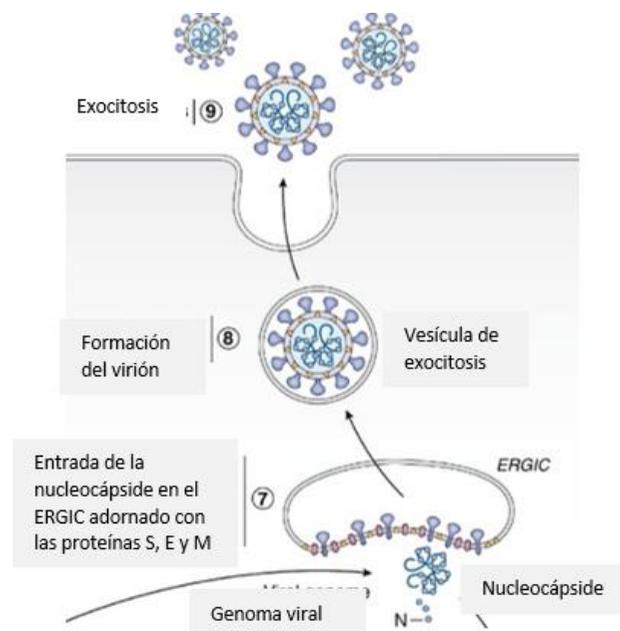


Figura 5 – Ensamblaje del virión y salida de la célula por exocitosis (modificado de Hartenian et al., 2020) – ERGIC (compartimento intermedio retículo endoplasmático-Golgi)

Cabe mencionar que para el correcto empaquetamiento del genoma en las nuevas partículas virales tiene que existir un mecanismo que reconozca qué ácidos nucleicos escoger, ya que durante la infección del virus se sintetizan tanto RNA positivos como negativos, además de los subgenómicos. En algunos coronavirus la señal de empaquetamiento se encuentra en la ORF1b (Sawicki, 2009).

### 4.3 Estructura genómica y expresión génica en Coronavirus

Los genomas de la familia *Coronaviridae* cuentan con un tamaño que oscila entre 27 y 32 kb ([Hartenian et al., 2020](#); [S. G. Sawicki & Sawicki, 1998](#)), con caperuza (5'-cap) en el extremo 5' y el extremo 3' poliadenilado ([Siddell et al., 1982](#)), siendo similares a los mRNAs de eucariotas. Son genomas RNA monohebra de sentido positivo no segmentados y policistrónicos, conteniendo varias ORFS que codifican para 22-29 proteínas. A su vez, ambos extremos del genoma serán regiones no traducidas o UTRs ([Sola et al., 2011](#)). El orden de los genes en este genoma de Coronavirus es 5'-replicasa-S-E-M-N-3', con proteínas accesorias intercaladas entre estas (Figura 6).



Figura 6 – Orden de los genes en algunos coronavirus (basado en Maroto-Vela & Piédrola-Angulo, 2020). L - Líder; PA - cola poli(A)

Cuentan con dos regiones codificadoras que ocupan los dos tercios del genoma más cercanos al extremo 5' (ORFs 1a y 1b) ([Nakagawa et al., 2020](#)). Serán procesadas cotraduccionalmente por proteasas virales, dando lugar a 15-16 proteínas no estructurales (nsps). La mayoría de estas nsps están implicadas en la síntesis del RNA viral. La replicasa viral será el único producto de la traducción derivado directamente del genoma. El resto de las proteínas codificadas por las ORFs del tercio más cercano al extremo 3', tanto estructurales como accesorias, se expresarán a partir de mRNA subgenómicos ([Sola et al., 2015](#)). Entre las accesorias encontraremos un número variable de proteínas dependiendo del coronavirus que, aunque no son esenciales para la replicación en cultivos celulares, parecen tener un papel en la patogenicidad de los virus ([Michel et al., 2020](#)). En la tabla 1 (anexo-I) se recogen las proteínas codificadas y su función.

#### 4.4 Traducción del genoma viral y cambio en el marco de lectura ribosómico (*ribosomal frameshifting*)

El cambio de marco de lectura ribosómico (*ribosomal frameshifting*) es un mecanismo de expresión génica usado por virus eucariotas (algunos retrovirus y varios virus de hebra RNA positiva entre otros) y bacteriófagos para expresar algunas de sus proteínas (Brierley & Dos Ramos, 2006).

La traducción del RNA viral da como primer producto una poliproteína, pp1a, codificada por la ORF1a. A continuación, tendrá lugar un cambio de marco de lectura ribosómico retrocediendo una posición (*backtrack*) -1 (poco antes del codón de terminación de la ORF1a), generando así la poliproteína 1ab (Plant & Dinman, 2008; Ziebuhr, 2004) (Figura 7). Posteriormente, ambas son procesadas por proteasas, dando lugar al complejo proteico replicasa-transcriptasa. Este mediará la replicación del genoma RNA viral y la transcripción de los mRNA subgenómicos, que a su vez codificarán las proteínas estructurales y accesorias en estos virus (Nakagawa et al., 2020).

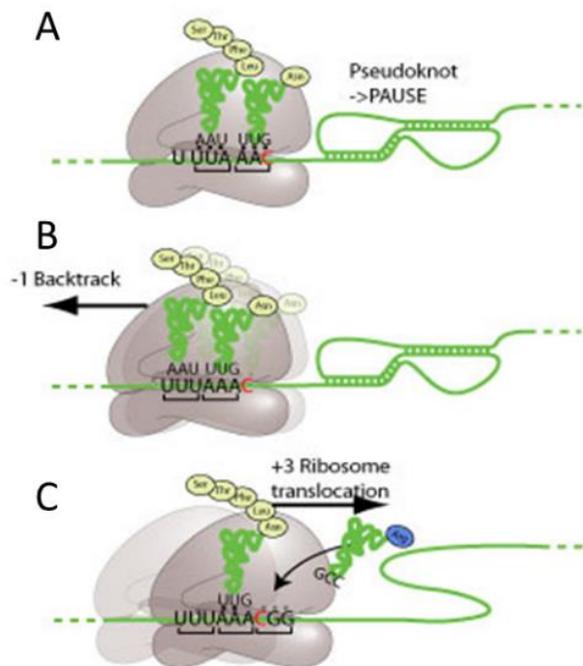


Figura 7 – *Ribosomal frameshifting* A – El ribosoma traduciendo alcanza la secuencia próxima al pseudonudo. B – Retroceso del ribosoma en una posición. C – Inicio de la traducción en nueva pauta de lectura. Secuencia deslizante en negrita (Modificada de *Ribosomal Frameshifting ~ ViralZone*, n.d.)

Este proceso de cambio de marco de lectura está dirigido por “señales específicas” presentes en el mRNA que fuerzan al ribosoma a cambiar el marco de lectura de la posición 0 al -1. Es decir, el ribosoma hará un movimiento en dirección 5', retrocediendo una posición y continuando la traducción en este nuevo marco. Las señales se encuentran en secuencias codificantes solapadas y, en la mayoría de los casos, el proceso es necesario para la expresión de replicasas (Hagemeijer et al., 2012).

En el caso de coronavirus de eucariotas, las “señales” para cambio de marco de lectura consisten en una “secuencia deslizante” y una estructura secundaria del RNA, con función estimuladora, que suele formar una estructura en pseudonudo (*pseudoknot*) (Firth & Brierley, 2012). Esta estructura (Figura 7) se encuentra 3' respecto a la secuencia deslizante; entre ambas se ubica una región espaciadora, de la cual se tiene que conservar cierta longitud para la máxima eficacia del proceso (Plant, 2008). Será en la secuencia deslizante donde el ribosoma cambie de marco de lectura y se considera que esto sucede por deslizamiento en tándem (*tandem slippage*) (Brierley & Dos Ramos, 2006).

La secuencia deslizante es un tramo de siete nucleótidos, con estructura XXXYYYYZ, en la que la X puede ser cualquier nucleótido, la Y será A o U y la Z tendrá que ser distinta a G. El deslizamiento en tándem tiene lugar en dos tRNAs anclados al ribosoma, que se deslizan desde la fase 0 (X\_XXY\_YYZ) a la fase -1 (XXX\_YYY). Se considera necesaria la naturaleza homopolimérica de esta secuencia, ya que así permite que el tRNA quede emparejado con el mRNA en, al menos, 2 de 3 posiciones de anticodón después del deslizamiento.

Gracias a este retroceso en la síntesis de proteínas se consiguen las poliproteínas 1a y 1ab.

Cabe destacar que la estequiometría de las proteínas codificadas por estas dos ORFs está conservada evolutivamente, por lo que podría ser necesaria para la replicación viral. El cambio en la pauta de lectura se encargaría de regular esta estequiometría, permitiendo producir el ratio requerido de las proteínas virales resultantes y regulando así los niveles de replicasas virales que, en elevadas cantidades, resultarían tóxicas (de Wilde et al., 2018).

#### **4.5 Proteínas de Coronavirus**

Tras obtener las dos poliproteínas, pp1a y pp1ab, estas sufrirán una escisión proteolítica, obteniéndose así 16 proteínas no estructurales, denominadas nsp1-nsp16, en SARS-CoV

(Ziebuhr, 2004). Derivadas de pp1a, encontraremos a las proteínas no estructurales (nsp) 1-11; a su vez, derivarán de pp1ab las nsp1-10 y nsp12-16. Nsp11 es un oligopéptido generado cuando no se da el cambio de marco de lectura. En la Tabla 1 (anexo-I) se incluye un sumario con las proteínas de estos virus así como sus funciones conocidas.

Los productos de la poliproteína 1a parecen preparar a la célula para la infección; aquellas exclusivas de pp1ab llevan a cabo la catálisis de la replicación y transcripción del RNA, junto a algunas enzimas celulares (de Wilde et al., 2018) y la proteína N (Robson et al., 2020). Los dominios transmembrana de algunas de estas proteínas no estructurales anclarán el complejo replicasa a membranas intracelulares, pudiendo estar involucradas en su remodelación y dando lugar a compartimentos de doble membrana dedicados a la síntesis del RNA viral (Masters, 2006).

#### **4.6 Replicación y transcripción en Coronavirus**

La síntesis de RNA RNA-dependiente en coronavirus engloba los procesos de replicación y de transcripción. En la replicación se sintetiza una hebra negativa complementaria a todo el genoma viral que posteriormente se empleará como molde para generar los genomas de la progenie (Sawicki, 2009; Hartenian 2020). Para que se inicie la síntesis de esta hebra molde, la RNA polimerasa RNA dependiente (RdRp) tiene que acceder al extremo 3' del genoma.

La transcripción en virus de esta familia tiene lugar en un proceso discontinuo (Sawicki et al., 2007) que implica un cambio de molde durante la síntesis de los RNA subgenómicos al que se añadirá una copia de la secuencia líder. Así se generarán varios RNA subgenómicos con ambos de sus extremos comunes con los del genoma viral además de la secuencia líder en su extremo 5' (Sola et al., 2015). Esta última secuencia presente únicamente en el extremo 5' del genoma sugiere que los RNAs subgenómicos son sintetizados mediante la fusión de secuencias no contiguas. De esta manera, se unirán a la secuencia líder y la secuencia codificante de cada RNA (también conocida como cuerpo o “body”) (Ziebuhr, 2004).

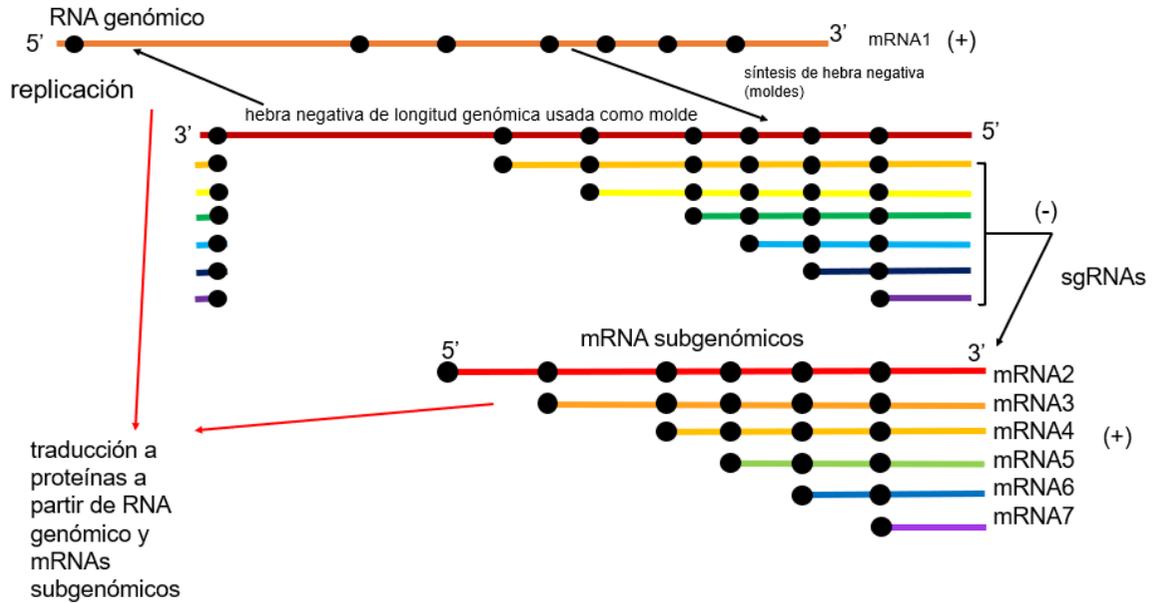


Figura 9 – RNAs genómicos y subgenómicos (basada en Sawicki, 2009). Los puntos indican las diferentes secuencias de regulación de la transcripción. sgRNA – RNAs subgenómicos

Este proceso de transcripción está controlado por las secuencias reguladoras de la transcripción (TRS), presentes en el extremo 3' de la secuencia líder (TRS-L) y antes de cada gen viral (TRS-B) (Fig 10A) (de Wilde et al., 2018; Stanley G. Sawicki, 2009; Sola et al., 2015). Estas TRS están formadas por una secuencia central (CS), de 6-7 nucleótidos, con regiones variables en cada uno de sus extremos (Dufour et al., 2011). La secuencia central es idéntica en el genoma (CS-L) y en los mRNAs (CS-B). Esto hace posible el apareamiento de la CS-L, del genoma, con la CS-B, de la hebra naciente negativa, llevando a la unión de la secuencia líder y la secuencia “body” (Sola et al., 2015).

La polimerasa reconocerá los TRS internos (TRS-L) y, tras copiarlos, se translocará al extremo 5' del genoma, a la secuencia líder (Snijder et al., 2016) (Fig 10B). Por tanto, la polimerasa RdRp puede unirse e iniciar la transcripción en diferentes puntos, continuando la elongación una vez copiado el gen de interés. Tras añadir la secuencia líder, se obtiene un molde de RNA subgenómico negativo. La transcriptasa viral reconocerá estos moldes y los copiará, generando mRNAs subgenómicos y genomas virales completos (Sawicki, 2009).

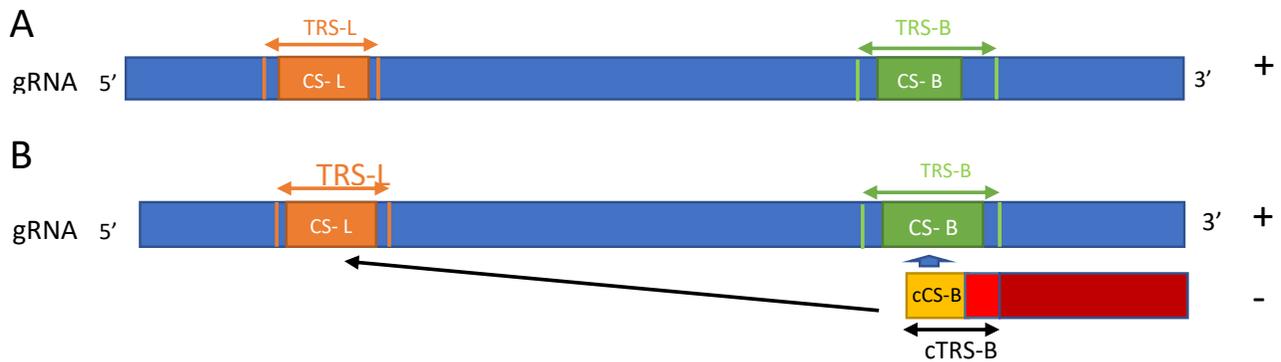


Figura 10 – A - Elementos de la secuencia involucrados en la transcripción discontinua. B - Representación del salto 3' a 5' en la transcripción discontinua durante la síntesis de hebras negativas. Basada en Zúñiga et al. (2004)

Estos saltos en la transcripción no sólo están regulados por las secuencias presentes en el RNA sino también por las estructuras secundarias en el RNA vírico, que a su vez forman una estructura tridimensional, y son esenciales para que se produzca este tipo de transcripción (Sola et al., 2015). Recientemente se han propuesto estas estructuras como diana terapéutica (Manfredonia et al., 2020).

Ya que los sgRNAs y el gRNA tienen el mismo extremo 5', compartiendo la señal de reconocimiento para el ribosoma, la abundancia relativa de las diferentes proteínas virales se determinará por la expresión relativa de sus RNAs (Snijder et al., 2016).

El modelo de transcripción que se da en Coronavirus incluye tres pasos. Para empezar, el RNA genómico formará precomplejos de iniciación, que acercan TRS-L y TRS-B. Interacciones entre RNA-RNA, proteína-RNA y proteína-proteína mantendrán estos precomplejos en equilibrio dinámico. Los precomplejos funcionarán como señales de freno y separación para el complejo de transcripción durante la síntesis de RNA de hebra negativa. Copiada la TRS-B, si la estabilidad ( $\Delta G$ ) del dúplex formado entre la hebra naciente negativa complementaria a TRS-B y TRS-L llega a cierto valor, ocurrirá un cambio de molde, añadiéndose una copia de la TRS-L y completándose así el RNA subgenómico de hebra negativa (Enjuanes et al., 2006; Sola et al., 2015).

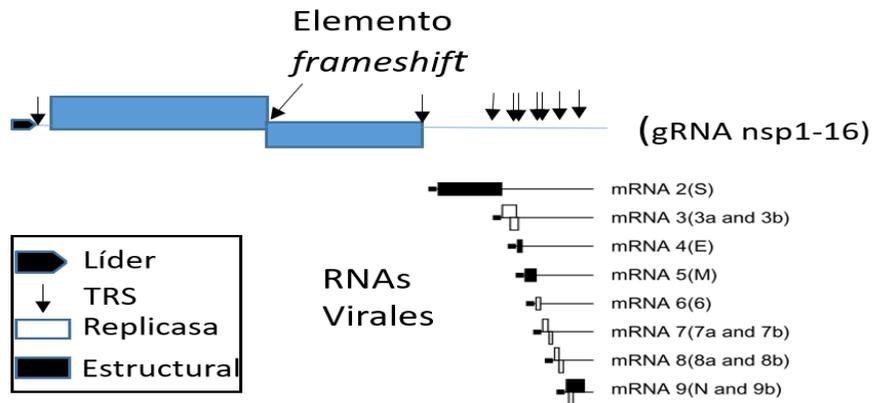


Figura 11 – Localización de las secuencias reguladoras de la transcripción (flechas verticales) en el genoma de coronavirus. (Modificada de Zúñiga et al., 2004)

Además de los productos RNA subgenómicos canónicos resultantes de este proceso de transcripción discontinua, se han encontrado en varios coronavirus productos RNA subgenómicos no canónicos, aunque se duda de si son resultados de este proceso o de la recombinación (V'kovski et al., 2020).

## 4.7 Técnicas para el estudio de coronavirus

### 4.7.1 Perfil ribosómico o *ribosome profiling*

Una de las metodologías para el estudio del cambio en la pauta de lectura de los ribosomas es el llamado “perfil ribosómico” o *ribosome profiling*, una técnica que se beneficia de la capacidad del ribosoma de proteger unos 30 nucleótidos del mRNA que está traduciendo, evitando su digestión por ribonucleasas (Ingolia et al., 2012). En este proceso, se lisan las células, tratándolas antes con inhibidores de la traducción y a continuación, con ribonucleasa eliminando así el RNA no protegido (Irigoyen et al., 2016). Posteriormente, tiene lugar la purificación de estos fragmentos de RNA, su secuenciación y mapeo (Brar &

Weissman, 2015; Ingolia et al., 2012). (Figura 12). Así, pueden localizarse y cuantificarse las regiones de RNA en los ribosomas con mucha precisión.

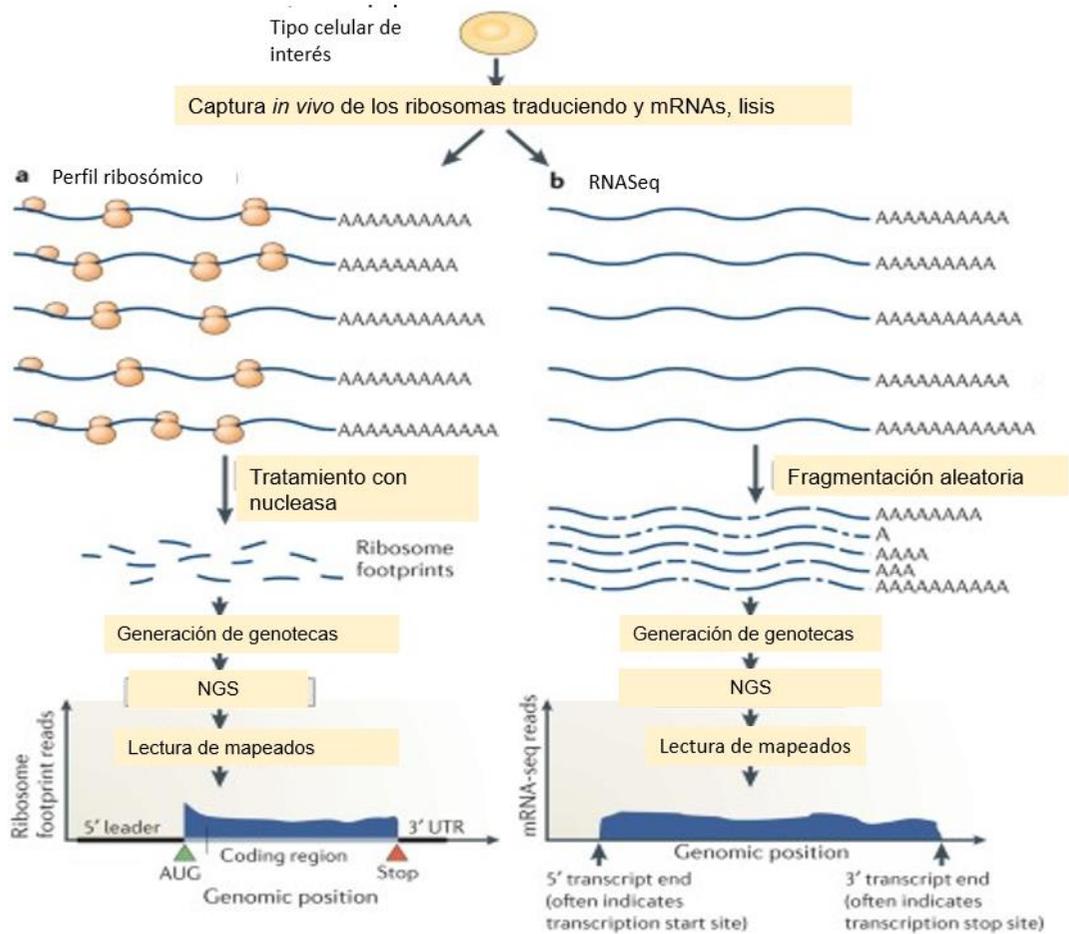


Figura 12 – Etapas de las metodologías de Ribosome profiling y RNASeq (modificado de Brar & Weissman, 2015).

#### 4.7.2 Secuenciado RNA y *ribosome profiling* en Coronavirus y ejemplos

El RNA-Seq es un método de secuenciación de RNA que usa el secuenciado empleando una metodología de alto rendimiento para obtener información del transcriptoma celular (Kukurba & Montgomery, 2015). Así, se pueden conocer las lecturas de genes presentes en la célula, como pueden ser el RNA ribosómico, tRNA, snRNA o small nuclear RNA y snoRNA o small nucleolar RNA, así como otras clases de RNAs no codificantes como los miRNAs y los piRNA, que regulan la expresión génica a nivel postranscripcional, y los lncRNAs (Kukurba & Montgomery, 2015). En un experimento típico de RNA-Seq, se sintetiza el cDNA

a partir del tipo de RNA de interés, se prepara la genoteca de secuenciado y se secuencia este cDNA en una plataforma NGS (Next Generation Sequencing) (Behjati & Tarpey, 2013).

Se pueden combinar las dos técnicas comentadas, como por ejemplo en el estudio de Irigoyen et. al. (2016) en el que se aplicaron las técnicas de RNA-Seq y *ribosome profiling* al MHV (*murine hepatitis virus*), infectando células de ratón. Las células se trataron con harringtonina (HAR) y/o cicloheximida (CHX) como inhibidores del inicio y la elongación de la traducción respectivamente. Posteriormente se extrajo el RNA. Se prepararon genotecas tanto para secuenciado RNA como para *ribosome profiling*. A continuación, se secuenciaron y mapearon los fragmentos de RNA, tanto a las secuencias virales como hospedadoras.

Una aproximación similar a otro coronavirus, en este caso el SARS-CoV-2, tuvo lugar en el estudio de Finkel et al. (2021).

Con estos estudios se caracterizó el genoma de estos virus y su expresión. También se analizó el ratio RiboSeq/RNASeq, que cuantifica la eficacia de la traducción (Blevins et al., 2018).

Así por ejemplo, Irigoyen y colaboradores (2016) observaron que la densidad ribosomal aumentaba de 5' a 3' cinco horas tras la infección, así como una reducción marcada en la densidad RiboSeq entre las ORF1a y el 1b, que mostraría los ribosomas que terminan en el codón de terminación de la primera ORF, no realizando el cambio de marco de lectura ribosómico. En cambio, la densidad de RNAs secuenciados fue constante entre estas dos primeras ORFs, aumentando luego de 5'>3'. Aun así, en la región 5' UTR se presenta una densidad RNASeq extra, que reflejará la presencia de la secuencia líder en todos los RNA, genómicos y subgenómicos. Esta se corresponde con una alta densidad de ribosomas unidos a RNA en esta secuencia líder 5', que no se corresponde con regiones codificantes conocidas; a su vez, la densidad de lecturas en la 3'-UTR es muy baja en comparación con las CDSs.

El RNA viral de sentido negativo se encontró en menores cantidades que el positivo en ambos estudios (Finkel et al., 2021; Irigoyen et al., 2016). En el segundo estudio, la ratio RiboSeq/RNASeq fue más baja en las ORF1ab que en el resto de las ORFs canónicas. Esto puede tener varias causas, o bien cierta proporción del RNA genómico no se traduce, estando en complejos de replicación-transcripción o bien está empaquetándose, o que alternativamente, tiene una eficacia de traducción relativamente baja.

Las eficacias de traducción de los RNAs virales está dentro del rango de los RNAs de la célula hospedadora. Esto indicaría que los transcritos virales no tienen preferencia durante la infección viral, si no que sus cantidades se deben a altos niveles de transcripción.

Además, este tipo de estudios permitió encontrar una ORF anterior a la ORF1a, que se considera regulador de la traducción del RNA genómico, la replicación viral y la patogénesis. Hay lecturas que mapean a este ORF, que se ha observado en varios CoVs. Así como se definió este, otras ORFs diferentes de las canónicas, han sido encontrados por métodos computacionales (Finkel et al., 2021).

El estudio de los virus de la familia *Coronaviridae*, resulta esencial en estos momentos debido a la actual pandemia causada por el SARS-CoV-2 y a la posibilidad de futuras pandemias víricas. Los mecanismos moleculares implicados en el ciclo viral, su replicación y la expresión de sus genes resultan, por tanto, un aspecto clave de cara a encontrar puntos esenciales para poder desarrollar tratamientos más efectivos para las enfermedades que estos causan. Por lo tanto, conocer los mecanismos citados, así como las proteínas generadas y sus funciones, es importante de cara a la prognosis y posterior tratamiento de las infecciones. Las investigaciones actuales contra el SARS-CoV-2 demuestran su auténtico potencial al haber permitido la obtención y desarrollo de vacunas en tiempo récord.

## 5. Conclusiones

En esta revisión hemos mostrado las peculiaridades de los virus de la familia *Coronaviridae*, su ciclo viral y su expresión génica. Caben destacar las siguientes conclusiones:

1. Los coronavirus tienen un genoma de un tamaño inusualmente grande. Esto probablemente responde al complejo mecanismo de regulación en estos virus por el cual se determina la expresión de sus proteínas.
2. La traducción de algunas de sus proteínas está controlada por el proceso de cambio de pauta de lectura *ribosomal frameshifting*, el cual ayudaría en la regulación de la estequiometría de las proteínas codificadas por las ORFs 1a y 1ab.
3. El resto de sus proteínas se sintetizan a partir de RNAs subgenómicos, generados por un proceso de transcripción discontinua.

4. Las secuencias reguladoras de la transcripción (TRS-L y TRS-B), encontradas tanto en la secuencia líder genómica como en las secuencias que codifican para proteínas, están encargadas de dirigir este proceso de transcripción discontinua.
5. Mediante la combinación de las técnicas de RiboSeq y RNASeq se puede medir la localización ribosómica en los RNAs víricos, la eficacia de la traducción de las proteínas, así como sus abundancias relativas. Gracias a la aplicación de estos procesos, junto con métodos computacionales, se pudo predecir y comprobar la expresión de nuevas ORFs. También han permitido comprobar la ausencia de pausado ribosomal en el sitio del *frameshifting*.

## 6. Bibliografía

- Behjati, S., & Tarpey, P. S. (2013). What is next generation sequencing? *Archives of Disease in Childhood: Education and Practice Edition*, 98, 236–238.
- Blevins, W., Tavella, T., Moro, S., Blasco-Moreno, B., Closa-Mosquera, A., Díez, J., Carey, L., & Albà, M. M. (2018). Using ribosome profiling to quantify differences in protein expression: a case study in *Saccharomyces cerevisiae* oxidative stress conditions. *BioRxiv*, 501478.
- Brar, G. A., & Weissman, J. S. (2015). Ribosome profiling reveals the what, when, where and how of protein synthesis. In *Nature Reviews Molecular Cell Biology*, 16, 651–664.
- Brierley, I., & Dos Ramos, F. J. (2006). Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Research*, 119, 29–42.
- Cornillez-Ty, C. T., Liao, L., Yates, J. R., Kuhn, P., & Buchmeier, M. J. (2009). Severe Acute Respiratory Syndrome Coronavirus Nonstructural Protein 2 Interacts with a Host Protein Complex Involved in Mitochondrial Biogenesis and Intracellular Signaling. *Journal of Virology*, 83, 10314–10318.
- de Wilde, A. H., Snijder, E. J., Kikkert, M., & van Hemert, M. J. (2018). Host factors in coronavirus replication. In *Current Topics in Microbiology and Immunology*, 419, 1–42.
- de Wit, E., van Doremalen, N., Falzarano, D., & Munster, V. J. (2016). SARS and MERS: recent insights into emerging coronaviruses. *Nature Reviews Microbiology*, 14, 523–534.

- Deng, X., & Baker, S. C. (2018). An “Old” protein with a new story: Coronavirus endoribonuclease is important for evading host antiviral defenses. *Virology*, *517*, 157–163.
- Dufour, D., Mateos-Gomez, P. A., Enjuanes, L., Gallego, J., & Sola, I. (2011). Structure and Functional Relevance of a Transcription-Regulating Sequence Involved in Coronavirus Discontinuous RNA Synthesis. *Journal of Virology*, *85*, 4963–4973.
- Egloff, M. P., Ferron, F., Campanacci, V., Longhi, S., Rancurel, C., Dutartre, H., Snijder, E. J., Gorbalenya, A. E., Cambillau, C., & Canard, B. (2004). The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 3792–3796.
- Enjuanes, L., Almazán, F., Sola, I., & Zuñiga, S. (2006). Biochemical aspects of coronavirus replication and virus-host interaction. In *Annual Review of Microbiology*, *60*, 211–230.
- Fehr, A. R., & Perlman, S. (2015). Coronaviruses: An overview of their replication and pathogenesis. In *Coronaviruses: Methods and Protocols*, 1282,1–23.
- Finkel, Y., Mizrahi, O., Nachshon, A., Weingarten-Gabbay, S., Morgenstern, D., Yahalom-Ronen, Y., Tamir, H., Achdout, H., Stein, D., Israeli, O., Beth-Din, A., Melamed, S., Weiss, S., Israely, T., Paran, N., Schwartz, M., & Stern-Ginossar, N. (2021). The coding capacity of SARS-CoV-2. *Nature*, *589*, 125–130.
- Firth, A. E., & Brierley, I. (2012). Non-canonical translation in RNA viruses. In *Journal of General Virology* *93*, 1385–1409.
- Gralinski, L. E., & Baric, R. S. (2015). Molecular pathology of emerging coronavirus infections. *Journal of Pathology*, *235*, 185–195.
- Hagemeyer, M. C., Rottier, P. J. M., & de Haan, C. A. M. (2012). Biogenesis and dynamics of the coronavirus replicative structures. In *Viruses*, *4*, 3245–3269.
- Hartenian, E., Nandakumar, D., Lari, A., Ly, M., Tucker, J. M., & Glaunsinger, B. A. (2020). The molecular virology of coronaviruses. *Journal of Biological Chemistry*, *295*(37), 12910–12934.
- Hulswit, R. J. G., de Haan, C. A. M., & Bosch, B. J. (2016). Coronavirus Spike Protein and Tropism Changes. In *Advances in Virus Research*, *96*, 29–57.

- Imbert, I., Guillemot, J.-C., Bourhis, J.-M., Bussetta, C., Coutard, B., Egloff, M.-P., Ferron, F., Gorbalenya, A. E., & Canard, B. (2006). A second, non-canonical RNA-dependent RNA polymerase in SARS Coronavirus. *The EMBO Journal*, 25, 4933–4942.
- Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M., & Weissman, J. S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nature Protocols*, 7, 1534–1550.
- Irigoyen, N., Firth, A. E., Jones, J. D., Chung, B. Y. W., Siddell, S. G., & Brierley, I. (2016). High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. *PLoS Pathogens*, Vol 12.
- Kukurba, K. R., & Montgomery, S. B. (2015). *Topic Introduction RNA Sequencing and Analysis*. 11, 951-969.
- Lai, M. M. C. (1990). Coronavirus: organization, replication and expression of genome. 44, 303-336
- Manfredonia, I., Nithin, C., Ponce-Salvatierra, A., Ghosh, P., Wirecki, T. K., Marinus, T., Ogando, N. S., Snijder, E. J., van Hemert, M. J., Bujnicki, J. M., & Incarnato, D. (2020). Genome-wide mapping of SARS-CoV-2 RNA structures identifies therapeutically-relevant elements. *Nucleic Acids Research*, 48, 12436–12452.
- Maroto-Vela, M. del C., & Piédrola-Angulo, G. (2020). Coronaviruses. *ANALES RANM*, 136, 235–238.
- Masters, P. S. (2006). The Molecular Biology of Coronaviruses. *Advances in Virus Research*, 65(January), 193–292.
- Menachery, V. D., Debbink, K., & Baric, R. S. (2014). Coronavirus non-structural protein 16: Evasion, attenuation, and possible treatments. In *Virus Research*, 194, 191–199.
- Michel, C. J., Mayer, C., Poch, O., & Thompson, J. D. (2020). Characterization of accessory genes in coronavirus genomes. *Virology Journal*, Vol 17
- Nakagawa, K., Lokugamage, K. G., & Makino, S. (2020). free information in English and Mandarin on the novel coronavirus COVID- Viral and Cellular mRNA Translation in Coronavirus- Infected Cells. *Advances in Virus Research*, 96(January), 165–192.
- Narayanan, K., Huang, C., Lokugamage, K., Kamitani, W., Ikegami, T., Tseng, C.-T. K., & Makino, S. (2008). Severe Acute Respiratory Syndrome Coronavirus nsp1 Suppresses

- Host Gene Expression, Including That of Type I Interferon, in Infected Cells. *Journal of Virology*, 82, 4471–4479.
- Neuman, B. W., Kiss, G., Kunding, A. H., Bhella, D., Baksh, M. F., Connelly, S., Droese, B., Klaus, J. P., Makino, S., Sawicki, S. G., Siddell, S. G., Stamou, D. G., Wilson, I. A., Kuhn, P., & Buchmeier, M. J. (2011). A structural analysis of M protein in coronavirus assembly and morphology. *Journal of Structural Biology*, 174, 11–22.
- Paules, C. I., Marston, H. D., & Fauci, A. S. (2020). Coronavirus Infections-More Than Just the Common Cold. In *JAMA - Journal of the American Medical Association*, 323, 707–708.
- Plant, E.P. (2008). The role of programmed-1 ribosomal frameshifting in coronavirus propagation. *Frontiers in Bioscience*, 13, 4873.
- Plant, E. P., & Dinman, J. D. (2008). The role of programmed-1 ribosomal frameshifting in coronavirus propagation. In *Frontiers in Bioscience*, 13, 4873–4881.
- Ribosomal frameshifting ~ ViralZone*. (n.d.). Retrieved March 17, 2021, from <https://viralzone.expasy.org/860>
- Robson, F., Shahed Khan, K., Le, T. K., Ment Paris, C., Demirbag, S., Barfuss, P., Rocchi, P., & Ng, W.-L. (2020). *Coronavirus RNA Proofreading: Molecular Basis and Therapeutic Targeting*. 710-727
- Sawicki, S. G., & Sawicki, D. L. (1998). A new model for coronavirus transcription. *Advances in Experimental Medicine and Biology*, 440, 215–219.
- Sawicki S.G. (2009) Coronavirus Genome Replication. En: Raney K., Gotte M., Cameron C. (eds) *Viral Genome Replication*. Springer, Boston, MA.
- Sawicki, S. G., Sawicki, D. L., & Siddell, S. G. (2007). A Contemporary View of Coronavirus Transcription. *Journal of Virology*, 81, 20–29.
- Schaecher, S. R., Mackenzie, J. M., & Pekosz, A. (2007). The ORF7b Protein of Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) Is Expressed in Virus-Infected Cells and Incorporated into SARS-CoV Particles. *Journal of Virology*, 81, 718–731.
- Schoeman, D., & Fielding, B. C. (2019). Coronavirus envelope protein: Current knowledge. In *Virology Journal*, 16, 1–22.
- Senapati, S., Banerjee, P., Bhagavatula, S., Kushwaha, P. P., & Kumar, S. (2021).

- Contributions of human ACE2 and TMPRSS2 in determining host–pathogen interaction of COVID-19. In *Journal of Genetics* (Vol. 100)
- Siddell, S., Wege, H., & ter Meulen, V. (1982). The structure and replication of coronaviruses. In *Current topics in microbiology and immunology*, 99, 131–163.
- Siu, Y. L., Teoh, K. T., Lo, J., Chan, C. M., Kien, F., Escriou, N., Tsao, S. W., Nicholls, J. M., Altmeyer, R., Peiris, J. S. M., Bruzzone, R., & Nal, B. (2008). The M, E, and N Structural Proteins of the Severe Acute Respiratory Syndrome Coronavirus Are Required for Efficient Assembly, Trafficking, and Release of Virus-Like Particles. *Journal of Virology*, 82, 11318–11330.
- Snijder, E. J., Decroly, E., & Ziebuhr, J. (2016). The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. In *Advances in Virus Research* 96, 59–126.
- Sola, I., Almazán, F., Zúñiga, S., & Enjuanes, L. (2015). Continuous and Discontinuous RNA Synthesis in Coronaviruses. In *Annual Review of Virology* (Vol. 2, pp. 265–288).
- Sola, I., Mateos-Gomez, P. A., Almazan, F., Zuñiga, S., & Enjuanes, L. (2011). RNA-RNA and RNA-protein interactions in coronavirus replication and transcription. *RNA Biology*, 8, 237–248.
- Sugiyama, K., & Amano, Y. (1981). Archives of Virology Morphological and Biochemical Properties of a New Coronavirus Associated with Diarrhea in Infant Mice. In *Archives of Virology*, 67, 241-251.
- Te Velhuis, A. J. W., Van Den Worm, S. H. E., & Snijder, E. J. (2012). The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Research*, 40, 1737–1747.
- V'kovski, P., Kratzel, A., Steiner, S., Stalder, H., & Thiel, V. (2020). Coronavirus biology and replication: implications for SARS-CoV-2. In *Nature Reviews Microbiology*, 19, 155–170.
- Verdiá-Báguena, C., Nieto-Torres, J. L., Alcaraz, A., DeDiego, M. L., Torres, J., Aguilera, V. M., & Enjuanes, L. (2012). Coronavirus E protein forms ion channels with functionally and structurally-involved membrane lipids. *Virology*, 432, 485–494.
- Yoshimoto, F. K. (2020). The Proteins of Severe Acute Respiratory Syndrome Coronavirus-

- 2 (SARS CoV-2 or n-COV19), the Cause of COVID-19. In *Protein Journal* 39, 198–216.
- Zhang, K., Miorin, L., Makio, T., Dehghan, I., Gao, S., Xie, Y., Zhong, H., Esparza, M., Kehrer, T., Kumar, A., Hobman, T. C., Ptak, C., Gao, B., Minna, J. D., Chen, Z., García-Sastre, A., Ren, Y., Wozniak, R. W., & Fontoura, B. M. A. (2021). Nsp1 protein of SARS-CoV-2 disrupts the mRNA export machinery to inhibit host gene expression. *Science Advances*, Vol 7.
- Ziebuhr, J. (2004). Molecular biology of severe acute respiratory syndrome coronavirus. *Current Opinion in Microbiology*, 7, 412–419.
- Zúñiga, S., Sola, I., Alonso, S., & Enjuanes, L. (2004). Sequence Motifs Involved in the Regulation of Discontinuous Coronavirus Subgenomic RNA Synthesis. *Journal of Virology*, 78, 980–994.

## Anexo- I

Tabla 1 - Proteínas codificadas por el genoma de SARS-CoV; (Modificada de Hartenian *et al.*, 2020).

<b>Tipos de proteínas</b>	<b>Nombre</b>	<b>Función</b>
<b>Proteínas estructurales</b>	<b>S</b>	Formación de peplómeros, unión a receptores para la entrada viral
	<b>E</b>	Proteína de envuelta, formación de viriones y ensamblaje
	<b>M</b>	Proteína de membrana, da forma al virión y curvatura a la membrana
	<b>N</b>	Proteína de la nucleocápside, unión a genoma RNA, asociación del genoma a RTCs, inhibición de la señalización celular y producción del interferón tipo 1 y supresión del silenciado RNA
<b>Proteínas no estructurales</b>	<b>nsp1</b>	Degradación del mRNA celular e inhibición de la traducción; inhibición del interferón tipo 1; preparación del medio celular para la replicación viral (Narayanan <i>et al.</i> , 2008; Zhang <i>et al.</i> , 2021)
	<b>nsp2</b>	Disruptor de la señalización intracelular (Cornillez-Ty <i>et al.</i> , 2009)
	<b>nsp3</b>	Proteasa de tipo papaína, escisión de polipéptidos, inhibición de la señalización y producción de interferones de tipo 1 (de Wilde <i>et al.</i> , 2018)
	<b>nsp4</b>	Formación de vesículas de doble membrana y fijación del complejo de replicación-transcripción a las membranas intracelulares (Masters, 2006)
	<b>nsp5</b>	Proteasa 3CL <sup>pro</sup> , escisión de polipéptidos, inhibición de la señalización de interferones de tipo I (Snijder <i>et al.</i> , 2016)
	<b>nsp6</b>	Restricción de la expansión de autofagosomas y formación de vesículas de doble membrana
	<b>nsp7</b>	Factor de procesividad para la RdRp y cofactor con nsp8 y nsp12

	<b>nsp8</b>	Factor de procesividad para la RdRp y cofactor con nsp7 y nsp12 (esto aumenta su afinidad al RNA (Te Velthuis et al., 2012)); actúa como RNA primasa, sintetizando oligonucleótidos cortos (Imbert et al., 2006; Stanley G. Sawicki, 2009)
	<b>nsp9</b>	Proteína de unión inespecífica al RNA (Egloff et al., 2004)
	<b>nsp10</b>	Aumento de la actividad catalítica de nsp14 y proteína de anclaje para nsp14 y nsp16
	<b>nsp11</b>	Desconocida (Yoshimoto, 2020)
	<b>nsp12</b>	RdRp
	<b>nsp13</b>	Helicasa de RNA, preparando el molde para nsp12 (Ziebuhr, 2004) actividad trifosfatasa
	<b>nsp14</b>	Exonucleasa de corrección (de Wit et al., 2016), N7-metiltransferasa
	<b>nsp15</b>	Endorribonucleasa (Deng & Baker, 2018), evasión de los sensores RNA de doble hebra
	<b>nsp16</b>	2'-O-metiltransferasa; añadiría capucha al RNA (Menachery et al., 2014)
<b>Proteínas accesorias</b>	<b>3a</b>	Formación de canales iónicos, inducción de necrosis, piroptosis, apoptosis y detención del ciclo celular
	<b>3b</b>	Inhibidor de la producción y señalización del interferón de tipo I, inductor de la apoptosis y necrosis, detención del ciclo celular
	<b>6</b>	Inhibición de la producción y señalización del interferón de tipo I y de la síntesis de DNA celular
	<b>7a</b>	Inhibición de la traducción en la célula hospedadora, inducción de la apoptosis y detención del ciclo celular
	<b>7b</b>	Componente del virión (Schaecher <i>et al.</i> , 2007)

	<b>8a</b>	Inducción de la apoptosis dependiente de caspasa
	<b>8b</b>	Aumento de la síntesis de DNA celular
	<b>9b</b>	Inducción de la apoptosis dependiente de caspasa