



Revista MINERVA

Plataforma digital de la revista: <https://minerva.sic.ues.edu.sv>



Origen del perfil de mutaciones presente en las secuencias de SARS-CoV-2 en El Salvador

Origin of the mutation profile present in SARS-CoV-2 sequences in El Salvador

Carlos Alexander Ortega Pérez¹
Noé Rigoberto Rivera²
Xochitl Sandoval López³
Carlos Enrique Hernández Ávila⁴

Correspondencia:
noe.rivera@ues.edu.sv

Presentado: 08 de diciembre de 2021
Aceptado: 23 de febrero de 2022

- 1 Departamento de Microbiología, Facultad de Medicina, Universidad de El Salvador.
orcid.org/0000-0003-1612-8386
- 2 Departamento de Bioquímica, Facultad de Medicina, Universidad de El Salvador.
orcid.org/0000-0001-9553-0554
- 3 Dirección, Instituto Nacional de Salud, El Salvador.
orcid.org/0000-0002-0988-1313
- 4 Departamento de Gobernanza, Instituto Nacional de Salud, El Salvador.
orcid.org/0000-0002-1230-0102

RESUMEN

Introducción: en el presente trabajo se describe el perfil de mutación y se analizan los distintos mecanismos responsables de las mutaciones en las primeras 6 secuencias completas del genoma de SARS-CoV-2 a partir de muestras de pacientes salvadoreños con diagnóstico de COVID-19. **Objetivo:** analizar el perfil de mutaciones de acuerdo a los mecanismos que dan origen a las mutaciones presentes en SARS-CoV-2. **Metodología:** se realizó un análisis de los cambios en las secuencias del genoma de SARS-CoV-2 utilizando como referencia la secuencia Wuhan (NC_045512.2), una vez conocidas las mutaciones, se procedió a tabular y generar gráficos de los SNPs y los genes afectados, además se analizaron los posibles mecanismos descritos responsables de generar las mutaciones estudiadas. **Resultados:** el análisis reveló que las mutaciones encontradas han sido reportadas a nivel mundial, sin embargo, las secuencias presentan mayor semejanza con los cambios descritos en Norteamérica, sumado a ello, el análisis global permitió clasificarlas en el caldo GISAID GH, y linaje pangolín B.1.2 y B.1.370, ambos linajes con una alta prevalencia en EUA, lo cual refuerza la hipótesis del origen norteamericano de las secuencias salvadoreñas. El patrón de cambios del genoma de SARS-CoV-2 en El Salvador, sugiere que las mutaciones son debidas a la acción de las desaminasas APOBEC (transición C>T) y ADARs (transición A>G), al efecto de especies reactivas de oxígeno (ROS) (transversión G>T), a errores propios del complejo replicación transcripción (RTC) que escapan a la corrección de la actividad exonucleasa de NSP14 y finalmente mutaciones como resultado de mecanismos de recombinación.

Palabras clave: SARS-CoV-2, D614G, NGS, 2019-nCoV, COVID-19s

ABSTRACT

Introduction: this paper describes the mutation profile and analyzes the different mechanisms responsible for mutations in the first 6 complete sequences of the SARS-CoV-2 genome from samples of Salvadoran patients diagnosed with COVID-19.

Objective: to analyze the mutation profile according to the mechanisms that give rise to the mutations present in SARS-CoV-2. **Methodology:** an analysis of the changes in the genome sequences of SARS-CoV-2 was performed using as reference the Wuhan sequence (NC_045512.2), once the mutations were known, we proceeded to tabulate and generate graphs of the SNPs and affected genes. The possible described mechanisms responsible for generating the mutations studied were also analyzed. **Results:** the analysis revealed that the mutations found have been reported worldwide, however, the sequences present greater similarity with the changes described in North America, added to this, the global analysis allowed classifying them in the GISAID GH broth, and pangolin lineage B.1.2 and B.1.370, both lineages with a high prevalence in the USA, which reinforces the hypothesis of the North American origin of the Salvadoran sequences. The pattern of changes in the SARS-CoV-2 genome in El Salvador suggests that the mutations are due to the action of APOBEC deaminase (C>T transition) and ADARs (A>G transition), to the effect of reactive oxygen species (ROS) (G>T transversion), to errors in the replication transcription complex (RTC) that escape the correction of the exonuclease activity of NSP14 and finally mutations as a result of recombination mechanisms.

Key words: SARS-CoV-2, D614G, NGS, 2019-nCoV, COVID-19s

INTRODUCCIÓN

El primer lugar en el que se identificó oficialmente el SARSCoV-2 fue en un mercado de mariscos en la ciudad de Wuhan, China, en el que existían las condiciones para la

transmisión de patógenos de animales salvajes a humanos (Zhu *et al.*, 2020) (Huang *et al.*, 2020). La secuenciación completa del genoma permitió clasificar a SARSCoV-2, como un nuevo virus dentro del género *Betacoronavirus*, de la subfamilia *Coronavirinae*, junto con SARS-CoV y MERS-CoV, con quienes comparte respectivamente el 79 y 50 % de similitud genómica (Gorbalenya *et al.*, 2020) (Lu *et al.*, 2020) (Zhou *et al.*, 2020). Inicialmente, la Organización Mundial de la Salud (OMS) lo denominó nuevo coronavirus 2019 (2019-nCoV) y, posteriormente, el Comité Internacional de Taxonomía de Virus lo denominó como Coronavirus 2 del Síndrome Respiratorio Agudo Grave (SARS-CoV-2) (Gorbalenya *et al.*, 2020).

Genoma de SARS-CoV-2

El genoma de SARS-CoV-2 está formado por una única cadena de ARN (ácido ribonucleico) monocatenario de polaridad positiva (+ssARN) de 29.82 kb, que codifica para 14 marcos de lectura abiertos. La cadena de ARN se asemeja, estructuralmente a un ARN mensajero (ARNm) de células eucarióticas, ya que, presenta un capuchón metilado en el extremo 5' (5' cap) y una cola poliadenilada (poli-A) en el extremo 3' (Chen *et al.*, 2020).

En el extremo 5' se encuentran los ORF1a y ORF1ab, estos se traducen en poliproteínas que se dividen en proteínas no estructurales (NSP) que forman el complejo replicación-transcripción (RTC) que consta de múltiples enzimas, incluida la proteasa similar a la papaína (NSP3), la proteasa principal (NSP5), el complejo primasa NSP7-NSP8, la ARN polimerasa dependiente de ARN de baja fidelidad o RdRp (NSP12), una helicasa trifosfatasa (NSP13), una exoribonucleasa (NSP14), una endonucleasa (NSP15), N7- y 2'-O- metiltransferasa (NSP10 y NSP16), respectivamente. En el extremo 3' otros ORFs codifican para al menos cuatro proteínas estructurales principales: proteína de espícula

(S), glicoproteína de membrana (M), envoltura (E) y nucleocápside (N) (Chen *et al.*, 2020) (Perlman & Netland, 2009)

El complejo replicación-transcripción

La replicación y la transcripción de los coronavirus ocurre en el citoplasma de las células infectadas y están mediados por el complejo replicación-transcripción (RTC). Dicho complejo se asocia con endosomas derivados del retículo endoplásmico rugoso que finalmente se convierten en “estructuras” virales de replicación (Snijder *et al.*, 2020). El anclaje del RTC a la membrana es mediado por NSP3 junto con proteínas transmembrana NSP4 y NSP6 (Angelini *et al.*, 2013). Las proteínas virales consideradas como el centro del complejo RTC son la ARN polimerasa dependiente de ARN (RdRp; gen NSP12), la helicasa NSP13 y los factores de procesividad NSP7 y NSP8 (Sola *et al.*, 2015).

Replicación y transcripción del genoma de SARS-CoV-2

Igual que otros virus ARN de hebra con sentido positivo, la replicación del genoma de los coronavirus incluye la síntesis de una hebra de ARN complementaria con sentido negativo que sirve de molde complementario de tamaño completo para la subsiguiente copia del templete en múltiples genomas con sentido positivo. La replicación está mediada por NSP12 o ARN polimerasa dependiente de ARN (RdRp) junto con otras enzimas replicativas codificadas por ORF1b que forman el complejo replicación-transcripción (RTC) (Subissi *et al.*, 2014). Existe evidencia que tanto el extremo 5´ como el 3´ del ARN genómico son indispensables para la producción de las hebras hijas de ARN con sentido positivo a partir del molde de ARN hebra negativa, lo anterior indica que la interacción de los extremos 5´ y 3´ del genoma

contribuyen a la replicación (Sola *et al.*, 2011). La replicación del ARN comienza con la unión de la RdRp al extremo 3´ de la hebra de ARN de sentido negativo, a estructuras secundarias de ARN y a secuencias específicas del genoma molde (Sola *et al.*, 2015).

Transcripción del genoma viral

La transcripción de la región proximal al extremo 3´ y los genes codificantes de proteínas accesorias es un proceso discontinuo que resulta en la producción de un set de mARN subgenómico (sgmARN). Los sgmARN de coronavirus poseen una secuencia líder en el extremo 5´ de cada genoma viral fusionado al extremo 5´ de cada mARN producido en la transcripción. Secuencias reguladoras de transcripción (TRSs) localizadas en el extremo 3´ de las secuencias líder (TRS; líder) y las TRS de cada gen (TRS-B; cuerpo) poseen secuencias complementarias que hibridan durante la transcripción y promueven la unión de las “TRS líder” y las “TRS cuerpo”, la presencia de las secuencias líder en cada transcrito protege a las moléculas de mARN viral de ser degradadas por la proteína NSP1 que promueve la degradación del mARN del hospedador (Sola *et al.*, 2015).

Mecanismos de corrección genética en coronavirus

Los coronavirus y otros del orden nidoviridae poseen los genomas ARN más grandes y complejos conocidos (Saberri *et al.*, 2018). La replicación de virus ARN típicamente presenta una alta tasa de error (o baja fidelidad) lo cual resulta en la existencia de una diversa población de genomas mutantes o “cuasiespecies” (Kuljić-Kapulica & Budisin, 1992). Si bien la baja fidelidad de la replicación permite a los virus ARN adaptarse y superar la presión de selección en diferentes entornos, también incrementa la posibilidad de errores

catastróficos que conlleven a la extinción viral. Lo anterior sugiere la necesidad de un equilibrio finamente ajustado entre la diversidad de las "cuasiespecies" y la tendencia hacia mutaciones que proporcionen mayor ventaja evolutiva y virulencia (Smith & Denison, 2012). La proteína NSP14 está altamente conservada entre los coronavirus, ya que su actividad exonucleasa correctora de errores es fundamental para la estabilidad y conservación de su complejo e inusualmente largo genoma (Gorbalenya *et al.*, 2006). La proteína NSP14 es una enzima bifuncional con un dominio N-terminal exonucleasa (ExoN) implicado en la fidelidad de la replicación y un dominio C-terminal con actividad N-metiltransferasa (N7-MTasa) implicado en la protección del mARN viral a través de la incorporación de un capuchón metilado en el extremo 5' (5' cap) del ARN (Chen *et al.*, 2009). Además, NSP14 está implicada en muchos procesos del ciclo de vida viral y patogenicidad, incluyendo la respuesta inmune innata y la recombinación del genoma viral (Becares *et al.*, 2016). El dominio N7-Metiltransferasa de NSP14 o dominio C-terminal N7-MTasa contiene un motivo de unión a DxG S-adenosil-L-metionina (SAM), este motivo es conservado entre los coronavirus y está implicado en la protección del mARN y la evasión de la respuesta inmune. Las mutaciones en el motivo de unión a DxG S-adenosil-L-metionina (SAM) de NSP14 de SARS-CoV-2 (mutación G333A) eliminan la actividad N7-MTasa afectan la replicación sin alterar la actividad ExoN (Case *et al.*, 2016).

La actividad exonucleasa de NSP 14 reside en el extremo o dominio N-terminal (Chen *et al.*, 2009). El dominio ExoN corrige los errores de la RdRp al remover los nucleótidos erróneamente apareados en el extremo 3' de la hebra de ARN en crecimiento (Ferron *et al.*, 2017). Además, reportes recientes demuestran que el dominio ExoN de NSP14 juega un rol en la frecuencia

de recombinación de sgRNA y el patrón de recombinación en células infectadas y viriones. La inactivación del dominio ExoN en NSP14 conlleva a una disminución de sgARN e incremento de genomas virales defectuosos (DVGs) (Gribble *et al.*, 2021).

Variación genética de SARS-CoV-2

Los virus son parásitos intracelulares estrictos, debido a ello los organismos han desarrollado mecanismos especializados en la detección y control viral (mecanismos de restricción viral). Entre los mecanismos de restricción viral encontramos la edición de ARN y ADN mediado por desaminasas endógenas que proveen una potente defensa contra virus específicos. Dos desaminasas están presentes en células de mamíferos: **ADARs** (ARN de doble cadena específicos de adenosina desaminasas) que actúan sobre ARN de doble cadena (dsARN) desaminando adeninas y generando inosina (A>I) que finalmente culmina en el cambio permanente A > G (Figura 1) (Eisenberg & Levanon, 2018) y **APOBEC** (Harris & Dudley, 2015) que desaminan citosinas a uracilo (C>U; C>T en ADN) en ácidos nucleicos de cadena sencilla (ssADN y ssARN) (Salter & Smith, 2018).

Durante la infección viral **ADARs** actúan directamente a través de la hipermutación del ARN viral o indirectamente a través de la edición de transcritos que modulan la respuesta celular (Taylor *et al.*, 2005) (Zahn *et al.*, 2007).

Se han reportado evidencias que ambos mecanismos de restricción viral (ADARs y APOBEC) están implicados en la edición del genoma de coronavirus, procesos que están implicados en la mutagénesis viral (Di Giorgio *et al.*, 2020). APOBECs actúa sobre el genoma viral, ya sea a través de la hipermutación C>U (Figura 2) o a través de una vía no enzimática que interfiere con la transcripción inversa (Pollpeter *et al.*, 2018).

Figura 1.

Generación de guanosina por acción de las desaminasas ADARs sobre el genoma viral.

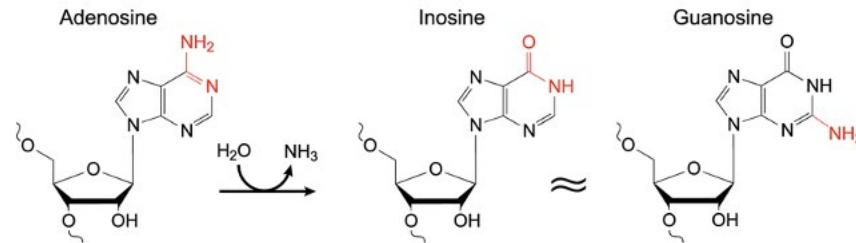
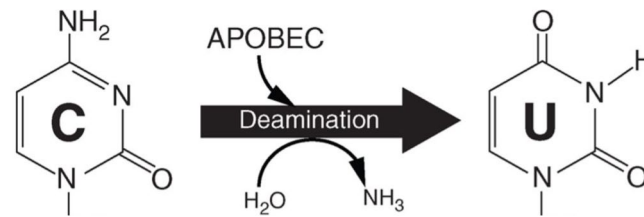


Figura 2.

Generación de uracilo por acción de las desaminasas APOBEC sobre el genoma viral.



Otro mecanismo indirectamente implicado en la mutagénesis es la producción de especies reactivas de oxígeno (ROS) relacionada con la actividad de APOBEC. ROS induce daño oxidativo de ácidos nucleicos transformando guanina en oxoguanina, está última se aparea con adenina y conlleva a la sustitución G> U (G>T en ADN) (Niocel *et al.*, 2019).

Un mecanismo no menos importante que los anteriores es la recombinación, que constituye una característica importante y particular del proceso evolutivo de los Coronavirus, resultados de algunos estudios han revelado una naturaleza en mosaico de la glicoproteína SARS-CoV-2 S que se le había restado importancia previamente, sin embargo, la evidencia indica que ocurrieron varios eventos de recombinación, involucrando linajes ancestrales Pangolin-CoV-2019, Bat-CoV-RaTG13 y Bat-CoV-RmYN02; a pesar de lo anterior, el muestreo limitado de Coronavirus en la vida silvestre puede complicar la inferencia definitiva del linaje parental de cada fragmento, incluso para segmentos genéticos como el RBD, donde existe una gran similitud entre Pangolin-CoV-2019 y SARS-CoV-2. Uno de estos eventos de recombinación condujo

a la adquisición de los residuos de unión a ACE2 que se encuentran en el linaje SARS-CoV-2 que circula actualmente en El Salvador, la nueva escisión similar a la furina en la unión S1/S2 también se encuentra dentro de uno de los posibles fragmentos de recombinación, esta característica de SARS-CoV-2 no estaba presente en los otros Coronavirus estrechamente relacionados o en los SARS-CoV que circulaban anteriormente (Tagliamonte *et al.*, 2021).

A la fecha han sido publicadas en GISAID 818,874 secuencias de genomas completos de SARS-CoV-2 (18/03/2021). A medida que más secuencias completas del genoma de SARS-CoV-2 se encuentran disponibles en las plataformas virtuales (GISAID, Nextstrain y otras), se va definiendo un patrón genético que permitirá establecer una relación directa entre el patrón de las mutaciones en el genoma de SARS-CoV-2 con regiones geográficas específicas (Mercatelli & Giorgi, 2020). El 22 de enero de 2021 se publican las primeras 6 secuencias del genoma de SARS-CoV-2 en El Salvador (números de Identificación GISAID EPI_ISL_672570, EPI_ISL_672572, EPI_ISL_672573, EPI_ISL_671974, EPI_ISL_671978 y

EPI_ISL_672012) (Hernandez Avila *et al.*, 2021).

Por otro lado, aún no está 100 % clara la base que explique las diferencias en virulencia, patogenicidad y mortalidad observados en diferentes países; todo indica que el análisis de más datos epidemiológicos y variación genómica permitirán conocer la razón de dichas diferencias (Brufsky, 2020). La información aportada por el análisis genómico, permitirá orientar el manejo farmacológico y eficacia de las vacunas (Wu *et al.*, 2020) y permitirá diseñar y aplicar estrategias de prevención y vigilancia epidemiológica más objetivas y eficientes (Zhao & Chen, 2020).

METODOLOGÍA

Para la determinación de las variantes por mutaciones con sentido erróneo y sinónimas se utilizó la plataforma bioinformática SOPHiA-DDMV5.7.10. En la verificación de las mutaciones, las secuencias fueron analizadas en la plataforma GISAID y Nexclade beta v0.8.1. Una vez obtenidos los informes de las mutaciones, se procedió a tabular y generar gráficos de los SNPs asociados al gen afectado y el respectivo cambio de aminoácido.

Consideraciones éticas

El protocolo de investigación fue sometido y aprobado por el Comité Local de Ética del INS; los investigadores no tuvieron acceso a la identidad de los pacientes u otros datos sensibles. Las muestras fueron manejadas por un número de referencia. Los datos de la secuenciación están disponibles en la plataforma GISAID bajo los siguientes registros: EPI_ISL_671974, EPI_ISL_671978, EPI_ISL_672012, EPI_ISL_672570, EPI_ISL_672572, EPI_ISL_672573.

RESULTADOS

Se analizaron las alteraciones genéticas

presentes en 6 secuencias del genoma completo de SARS-CoV-2 reportadas en El Salvador (Hernandez Avila *et al.*, 2021), las cuales están disponibles en GISAID.

Los resultados reflejan un total de 48 mutaciones diferentes al compararlas con el genoma de referencia Wuhan (NC_045512.2). Entre las más frecuentes, 7 mutaciones están presente en todas las secuencias analizadas (A23403G, C14408T, C3037T, C241T, G25563T, C1059T, C27964T) y 4 están presente en 5 secuencias (C26164T, G13617T, C27509T, G28899A), el resto de mutaciones **únicamente** está presente en una secuencia (Figura 3).

7 mutaciones están presente en todas las secuencias analizadas (A23403G, C14408T, C3037T, C241T, G25563T, C1059T, C27964T) y 4 están presente en 5 secuencias (C26164T, G13617T, C27509T, G28899A), otras mutaciones se encuentran en 2 y otras en 1 secuencia.

Únicamente se han encontrado cambios debido a polimorfismo de nucleótido simple (SNPs); no se evidenciaron inserciones ni deleciones. El 56.25% (27 diferentes mutaciones) provocan cambios en la secuencia de las proteínas (mutaciones con sentido erróneo), el 44.75 % (21 mutaciones) no generan cambios (mutaciones silenciosas) y el 2.08 % (1 mutación) se encuentran en la región reguladora 5´UTR (Figura 4).

En las secuencias analizadas, se detectaron 48 mutaciones, de las cuales el 56.25 % (27 cambios) son SNPs con sentido erróneo, 43.75 % (21 cambios) son mutaciones silenciosas. El 2.05 % ocurren en la región 5´UTR, no se encontraron mutaciones en el extremo 3´UTR, deleción ni inserciones. Las diferentes secuencias presentan un rango entre 15 y 19 mutaciones, mostrando un promedio de 16.7 mutaciones por secuencia (Figura 5).

Se reportan entre 15 y 19 mutaciones por aislado. El promedio de mutaciones por aislado es 16.7.

Figura 3.

Mutaciones reportadas a nivel de ácidos nucleicos en el geoma de SARS-CoV-2 en El Salvador.

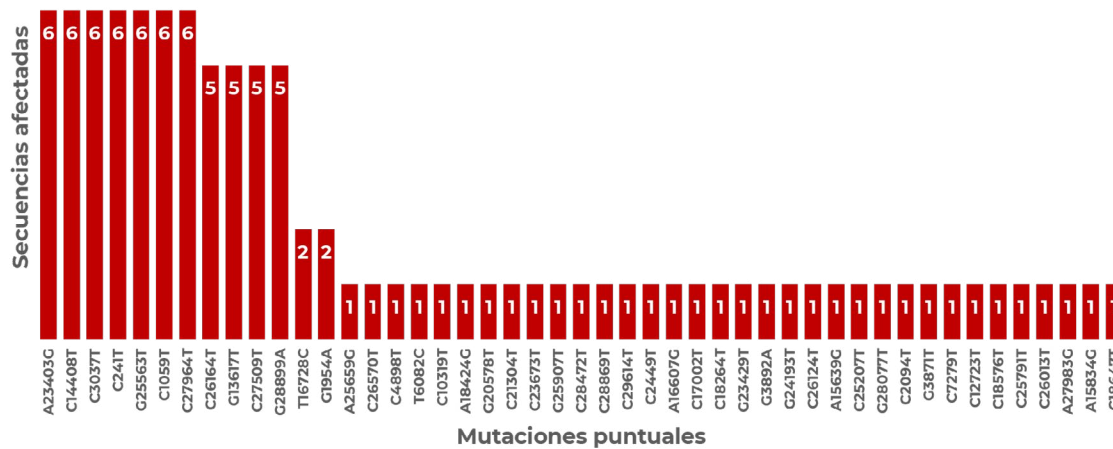


Figura 4.

Tipos de mutaciones encontradas en las secuencias de SARS-CoV-2 de El Salvador.

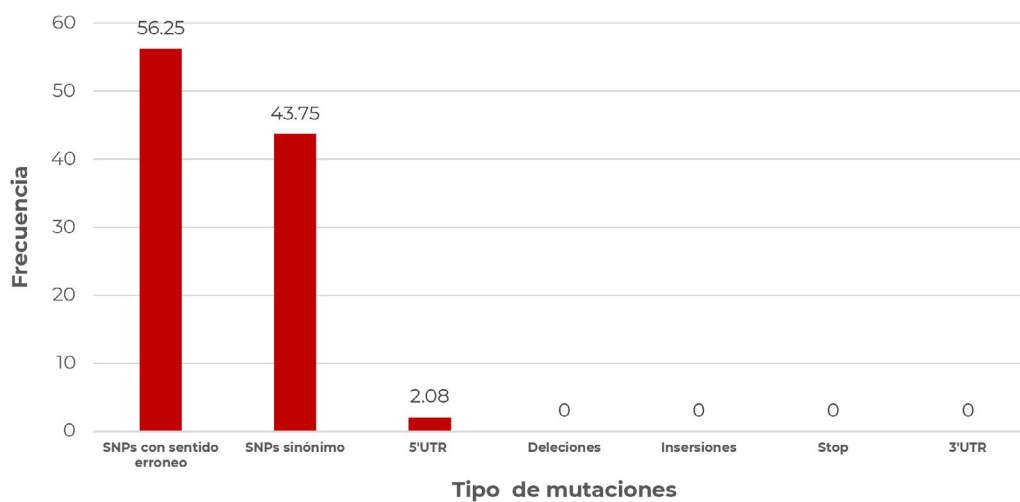
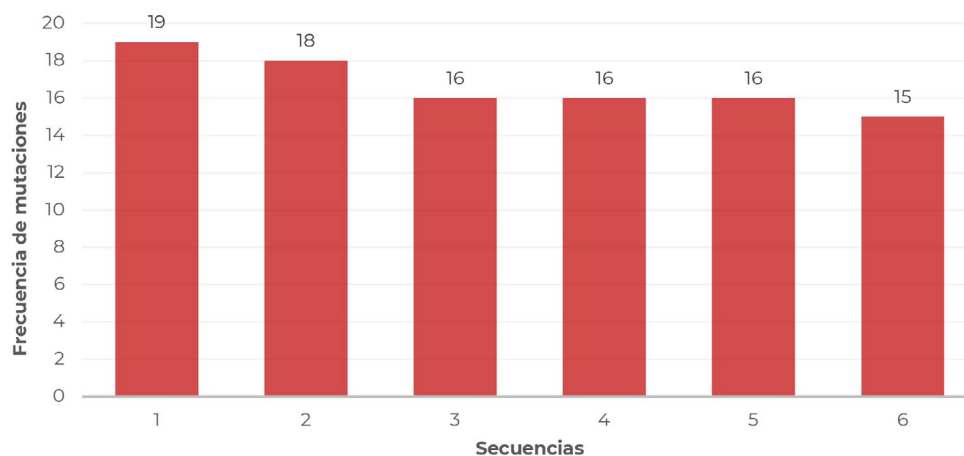


Figura 5.

Número de mutaciones por secuencia.

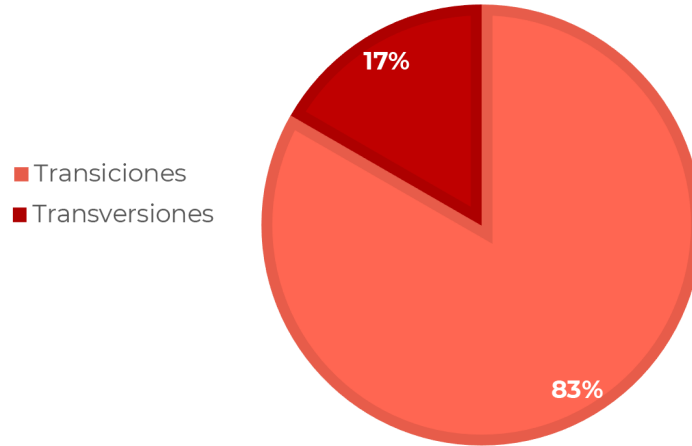


Las transiciones representan el 83 % del total de mutaciones encontradas, mientras que las

transversiones el 17 %. La única transversión encontrada es la sustitución G>T (Figura 6).

Figura 6.

Tipos de sustituciones presentes en las secuencias analizadas.



Predominan las transiciones (83 %) sobre transversiones (17 %).

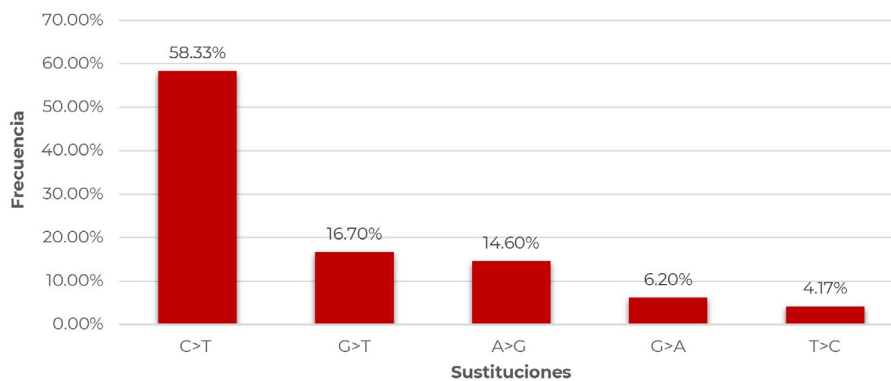
De mayor a menor porcentaje las sustituciones encontradas son C>T (58.3 %), G>T (16.7 %), A>G (14.6 %), G>A (6,2 %) y T>C (4.17 %) (Figura 7).

Los genes afectados con mayor número de mutaciones con sentido erróneo (3 cambios) son, **S** (D614G, A623S y S704L), **NS3** (P258S, Q57H y G172V), **NSP2** (S430L, T85I, T265) y N (N R209K, N P67S, NP199L); \con 2 mutaciones los genes NS8, NSP12, NSP3 y NSP13 y con una mutación los genes NSP5, NSP9, NSP15 y NS7a (Figura 8).

Entre las mutaciones con sentido erróneo 6 (S^{D614G}, NS8^{S24L}, NS3^{Q57H}, NSP2^{T85I}, NSP2^{T265I}, NSP12^{P323L}) están presentes en el 100 % de las secuencias (6 de 6 aislados); 4 mutaciones (NS7a^{T39I}, NSP12^{K59N}, N^{R209K} y NS3^{P258S}) están presentes en el 83.3 % de las muestras (5 de 6) y otras mutaciones se encuentran únicamente en una (Figura 9). En ninguna de las secuencias se observaron mutaciones propias de las variantes del Reino Unido, Sudáfrica, Brasil y California que presentan mayor transmisibilidad e infectividad.

Figura 7.

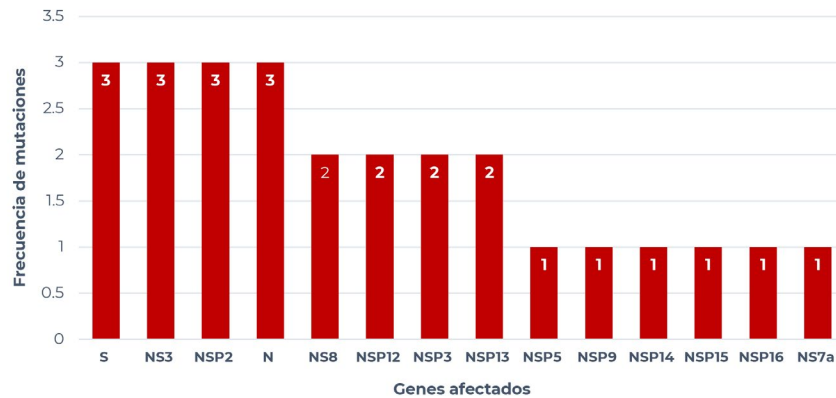
Tipos de sustituciones encontradas en los genomas analizados.



Nota. Existe un predominio de transiciones sobre transversiones. Entre las transiciones el cambio C>T representa el 58.33 % de todos los cambios encontrados. Únicamente se encontró la transversión G>T.

Figura 8.

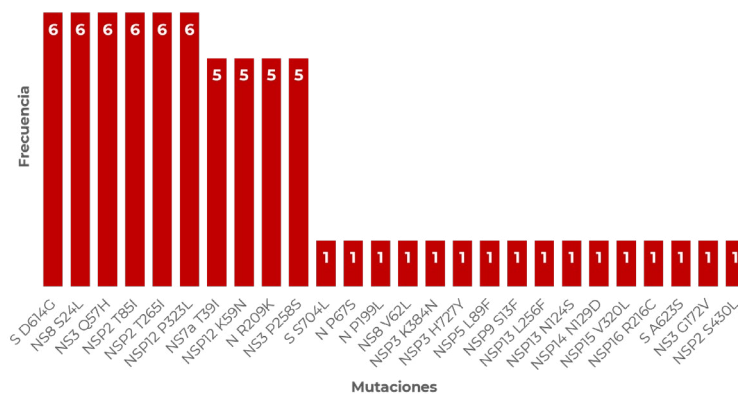
Genes afectados por mutaciones.



Nota. Los genes afectados con mayor número de mutaciones con sentido erróneo son el gen **S** (D614G, A623S y S704L), **NS3** (P258S, Q57H y G172V) y **NSP2** (S430L, T85I, T265I) y N (R209K, P67S y P199L). Con 2 mutaciones los genes NS8, NSP12, NSP3 y NSP13 y con una mutación los genes NSP5, NSP9, NSP15 Y NS7a.

Figura 9.

Frecuencia de mutaciones con sentido erróneo encontradas en las secuencias publicadas en El Salvador.



Nota. 6 mutaciones (S D614G, NS8 S24L, NS3 Q57H, NSP2 T85I, NSP2 T265I y NSP12 P323L) están presente en el 100 % de los aislados (6 de 6 aislados); 4 mutaciones (NS7a T39I, NSP12 K59N, N R209K y NS3 P258S) están presente en el 83.3 % de las muestras (5 de 6) y otras mutaciones se encuentran únicamente en una muestra.

DISCUSIÓN

A nivel mundial se ha encontrado un promedio de 7.23 mutaciones por secuencia (Mercatelli & Giorgi, 2020); en El Salvador el promedio es de 16.7 mutaciones por secuencia. En relación a las mutaciones en regiones codificantes, es semejante a lo reportado a nivel global, ya que en El Salvador predominan los SNPs con sentido erróneo (55.3 %) sobre las mutaciones sinónimas (44.68 %). En El Salvador las regiones reguladoras presentan la mutación 5'UTR 241C>T, dicha mutación también está presente

en el 100 % de las secuencias analizadas a nivel mundial. Actualmente, se está acumulando evidencia sobre el efecto de la mutación 5'UTR 241C>T sobre la replicación, la transcripción y el plegamiento de los ssRNA virales, lo cual continúa en estudio (Sola *et al.*, 2011) and the transcription of sgRNAs by a discontinuous mechanism. Both processes are regulated by RNA sequences such as the 5' and 3' untranslated regions (UTRs. En las 6 secuencias analizadas no se evidencia ninguna deleción, inserción, o codones de parada reportadas en otras regiones.

El perfil de sustitución de El Salvador es semejante al encontrado a nivel mundial; muestra un predominio de las transiciones (Alouane *et al.*, 2020), en particular las transiciones C>T (58.3 %) y A>G (14.6 %). En el caso de las transversiones, solo se detectó el cambio G>T (16.7 %).

Este patrón de cambios en el genoma de SARS-CoV-2 en El Salvador, sugiere que el origen de las mutaciones observadas en las secuencias de El Salvador incluye los errores propios del complejo replicación transcripción (RTC) que escapan a la corrección de la actividad exonucleasa de NSP14 (Ferron *et al.*, 2017) y mayoritariamente a la participación de los mecanismos de restricción viral del sistema inmune humano, mediado por las desaminasas APOBEC (Harris & Dudley, 2015) que desaminan citosinas a uracilo y se refleja en el ADN como el cambio C>T, ADARs que desamina adeninas y finalmente culmina en el cambio permanente A > G (Eisenberg & Levanon, 2018) y a la producción de especies reactivas de oxígeno (ROS) relacionada con la actividad de APOBEC, la cual induce daño oxidativo de ácidos nucleicos que conlleva a la transversión G> U (G>T en ADN) (Niocel *et al.*, 2019).

A pesar que la mayoría de las mutaciones encontradas en El Salvador han sido reportadas en todo el mundo, la proporción de cada uno de los cambios a nivel de ácidos nucleicos, los genes más afectados y las mutaciones con sentido erróneo, indica que el patrón de mutaciones en las secuencias de El Salvador tiene mayor semejanza con las secuencias publicadas en Norteamérica, lo que sugiere que fueron importadas de dicha región. Por otro lado, al analizar las 6 secuencias en las plataformas virtuales Nextstrain y GISAID todas fueron clasificadas en el clado Nextstrain 20C, caracterizado por la mutación A23403G, C1059T y G25563T y según GISAID pertenecen al clado

GH, linaje pangolín B.1.2 y B.1.370 (Hernandez Avila *et al.*, 2021) , ambos linajes con una alta prevalencia en EUA.

CONCLUSIONES

Podemos asegurar que las mutaciones encontradas en El Salvador coinciden con mutaciones que ya habían sido reportadas a nivel mundial y por tanto no se pudo determinar si algunas se originaron de manera autóctona. Las secuencias presentan mayor semejanza con los cambios descritos en Norteamérica, sumado a ello, el análisis global permitió clasificarlas en el caldo GISAID GH, y linaje pangolín B.1.2 y B.1.370, ambos linajes con una alta prevalencia en EUA, lo cual refuerza la hipótesis del origen norteamericano de las secuencias salvadoreñas. El patrón de cambios del genoma de SARS-CoV-2 en El Salvador, sugiere que las mutaciones se deben a la acción de las desaminasas APOBEC (transición C>T) y ADARs (transición A>G), al efecto de especies reactivas de oxígeno (ROS) (transversión G>T), a errores propios del complejo replicación transcripción (RTC) que escapan a la corrección de la actividad exonucleasa de NSP14 y finalmente mutaciones como resultado de mecanismos de recombinación. Respecto a los eventos de recombinación que conducen a la adquisición de los residuos de unión a ACE2 que se encuentran en algunos linajes de SARS-CoV-2 que circulan a nivel mundial y en nuestro país, en donde la escisión similar a la furina en la unión S1 / S2 también se encuentra dentro de uno de los posibles fragmentos afectados, será necesario más estudio y análisis del fenómeno de recombinación en SARS-CoV2 de secuencias de El Salvador que se dará en una próxima entrega.

Es necesario continuar secuenciando más genomas de SARS-CoV-2 aislados de pacientes de El Salvador, con el objetivo de verificar si se detectan patrones de mutaciones típicos

de otras regiones geográficas; tales como las variantes inglesa, sudafricana, brasileña y de California; además, vigilar el apareamiento de mutaciones no reportadas en otros países y analizar sus posibles efectos sobre la infectividad y transmisibilidad, la eficacia farmacológica (antivirales y vacunas) y el efecto en las regiones utilizadas para la detección y diagnóstico.

CONFLICTO DE INTERÉS Y FINANCIAMIENTO

Los autores certifican que no existe ningún conflicto de interés que afecte el presente trabajo. Todos los experimentos se realizaron con equipos y reactivos de la Universidad de El Salvador en el marco del proyecto de secuenciación de SARS-CoV2 en El Salvador y el proyecto 15-04 que se realizan actualmente en la Sección de Virología y Microbiología Molecular del Departamento de Microbiología de la Facultad de Medicina, Universidad de El Salvador.

Todos los experimentos para obtener las secuencias de SARS-CoV2 se realizaron en las instalaciones de la Sección de Virología y Microbiología Molecular del Departamento de Microbiología de la Facultad de Medicina, Universidad de El Salvador.

AGRADECIMIENTOS

Al Departamento de Microbiología de la Facultad de Medicina por contribuir a generar las condiciones en Laboratorio de Virología para la realización de este trabajo. A las autoridades de la Universidad de El Salvador y de la Facultad de Medicina por el apoyo logístico y demás recursos de laboratorio necesarios para la realización de la presente investigación. Al Consejo de Investigaciones

Científicas de la Universidad de El Salvador (CIC-UES) en particular por el apoyo decidido a la investigación científica en El Salvador.

REFERENCIAS BIBLIOGRÁFICAS

Alouane, T., Laamarti, M., Essabbar, A., Hakmi, M., Bendani, H., Laamarti, R., Ghrifi, F., Allam, L., Aanniz, T., Mentag, R., Sbabou, L., Nejari, C., Amzazi, S., & Belyamani, L. (2020). *SARS-CoV-2 Genomes: Moving Toward a Universal Vaccine for the "Confined Virus"?* 1-19.

Angelini, M. M., Akhlaghpour, M., Neuman, B. W., & Buchmeier, M. J. (2013). Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. *MBio*, 4(4), 1-10. <https://doi.org/10.1128/mBio.00524-13>

Becares, M., Pascual-iglesias, A., Nogales, A., Sola, I., Enjuanes, L., & Zuñiga, S. (2016). Mutagenesis of Coronavirus nsp14 Reveals Its Potential Role in. *Journal of Virology*, 90(11), 5399-5414. <https://doi.org/10.1128/JVI.03259-15>. Editor

Brufsky, A. (2020). Distinct viral clades of SARS-CoV-2: Implications for modeling of viral spread. *Journal of Medical Virology*, 92(9), 1386-1390. <https://doi.org/10.1002/jmv.25902>

Case, J. B., Ashbrook, A. W., Dermody, T. S., & Denison, M. R. (2016). Mutagenesis of S -Adenosyl-I-Methionine-Binding Residues in Coronavirus nsp14 N7-Methyltransferase Demonstrates Differing Requirements for Genome Translation and Resistance to Innate Immunity. *Journal of Virology*, 90(16), 7248-7256. <https://doi.org/10.1128/jvi.00542-16>

Chen, Y., Cai, H., Pan, J., Xiang, N., Tien, P.,

- Ahola, T., & Guo, D. (2009). Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(9), 3484–3489. <https://doi.org/10.1073/pnas.0808790106>
- Chen, Y., Liu, Q., & Guo, D. (2020). Emerging coronaviruses: Genome structure, replication, and pathogenesis. *Journal of Medical Virology*, *92*(4), 418–423. <https://doi.org/10.1002/jmv.25681>
- Di Giorgio, S., Martignano, F., Torcia, M. G., Mattiuz, G., & Conticello, S. G. (2020). Evidence for host-dependent RNA editing in the transcriptome of SARS-CoV-2. *BioRxiv*, June, 1–8. <https://doi.org/10.1101/2020.03.02.973255>
- Eisenberg, E., & Levanon, E. Y. (2018). A-to-I RNA editing—Immune protector and transcriptome diversifier. *Nature Reviews Genetics*, *19*(8), 473–490. <https://doi.org/10.1038/s41576-018-0006-1>
- Ferron, F., Subissi, L., De Morais, A. T. S., Le, N. T. T., Sevajol, M., Gluais, L., Decroly, E., Vornrhein, C., Bricogne, G., Canard, B., & Imbert, I. (2017). Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(2), E162–E171. <https://doi.org/10.1073/pnas.1718806115>
- Gorbalenya, A. E., Baker, S. C., Baric, R. S., de Groot, R. J., Drosten, C., Gulyaeva, A. A., Haagmans, B. L., Lauber, C., Leontovich, A. M., Neuman, B. W., Penzar, D., Perlman, S., Poon, L. L. M., Samborskiy, D. V., Sidorov, I. A., Sola, I., & Ziebuhr, J. (2020). The species Severe acute respiratory syndrome-related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology*, *5*(4), 536–544. <https://doi.org/10.1038/s41564-020-0695-z>
- Gorbalenya, A. E., Enjuanes, L., Ziebuhr, J., & Snijder, E. J. (2006). Nidovirales: Evolving the largest RNA virus genome. *Virus Research*, *117*(1), 17–37. <https://doi.org/10.1016/j.virusres.2006.01.017>
- Gribble, J., Stevens, L. J., Agostini, M. L., Anderson-Daniels, J., Chappell, J. D., Lu, X., Pruijssers, A. J., Routh, A. L., & Denison, M. R. (2021). The coronavirus proofreading exoribonuclease mediates extensive viral recombination. *PLoS Pathogens*, *17*(1), 1–28. <https://doi.org/10.1371/journal.ppat.1009226>
- Harris, R. S., & Dudley, J. P. (2015). APOBECs and virus restriction. *Virology*, *479*–*480*, 131–145. <https://doi.org/10.1016/j.virol.2015.03.012>
- Hernandez Avila, C. E., Ortega Perez, C. A., Rivera, N. R., & López, X. S. (2021). Primeras seis secuencias del genoma completo de SARS-CoV-2 por NGS en El Salvador. *Alerta, Revista Científica Del Instituto Nacional de Salud*, *4*(1), 61–66. <https://doi.org/10.5377/alerta.v4i1.10682>
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., ... Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, *395*(10223), 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
- Kuljić-Kapulica, N., & Budisin, A. (1992). Coronaviruses. *Srpski Arhiv Za Celokupno Lekarstvo*, *120*(7–8), 215–218. <https://doi.org/10.4161/rna.8.2.15013>
- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B.,

- Zhu, N., Bi, Y., Ma, X., Zhan, F., Wang, L., Hu, T., Zhou, H., Hu, Z., Zhou, W., Zhao, L., ... Tan, W. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *The Lancet*, 395(10224), 565–574. [https://doi.org/10.1016/S0140-6736\(20\)30251-8](https://doi.org/10.1016/S0140-6736(20)30251-8)
- Mercatelli, D., & Giorgi, F. M. (2020). Geographic and Genomic Distribution of SARS-CoV-2 Mutations. *Frontiers in Microbiology*, 11(July), 1–13. <https://doi.org/10.3389/fmicb.2020.01800>
- Niocol, M., Appourchaux, R., Nguyen, X. N., Delpeuch, M., & Cimarelli, A. (2019). The DNA damage induced by the Cytosine Deaminase APOBEC3A Leads to the production of ROS. *Scientific Reports*, 9(1), 1–11. <https://doi.org/10.1038/s41598-019-40941-8>
- Perlman, S., & Netland, J. (2009). Coronaviruses post-SARS: Update on replication and pathogenesis. *Nature Reviews Microbiology*, 7(6), 439–450. <https://doi.org/10.1038/nrmicro2147>
- Phuphuakrat, A., Kraiwong, R., Boonarkart, C., Lauhakirti, D., Lee, T.-H., & Auewarakul, P. (2008). Double-Stranded RNA Adenosine Deaminases Enhance Expression of Human Immunodeficiency Virus Type 1 Proteins. *Journal of Virology*, 82(21), 10864–10872. <https://doi.org/10.1128/jvi.00238-08>
- Pollpeter, D., Parsons, M., Sobala, A. E., Coxhead, S., Lang, R. D., Bruns, A. M., Papaioannou, S., McDonnell, J. M., Apolonia, L., Chowdhury, J. A., Horvath, C. M., & Malim, M. H. (2018). Deep sequencing of HIV-1 reverse transcripts reveals the multifaceted antiviral functions of APOBEC3G. *Nature Microbiology*, 3(2), 220–233. <https://doi.org/10.1038/s41564-017-0063-9>
- Saber, A., Gulyaeva, A. A., Brubacher, J. L., Newmark, P. A., & Gorbalenya, A. E. (2018). A planarian nidovirus expands the limits of RNA genome size. In *PLoS Pathogens* (Vol. 14, Issue 11). <https://doi.org/10.1371/journal.ppat.1007314>
- Salter, J. D., & Smith, H. C. (2018). Modeling the Embrace of a Mutator: APOBEC Selection of Nucleic Acid Ligands. *Trends in Biochemical Sciences*, 43(8), 606–622. <https://doi.org/10.1016/j.tibs.2018.04.013>
- Smith, E. C., & Denison, M. R. (2012). Implications of altered replication fidelity on the evolution and pathogenesis of coronaviruses. *Current Opinion in Virology*, 2(5), 519–524. <https://doi.org/10.1016/j.coviro.2012.07.005>
- Snijder, E. J., Limpens, R. W. A. L., de Wilde, A. H., de Jong, A. W. M., Zevenhoven-Dobbe, J. C., Maier, H. J., Faas, F. F. G. A., Koster, A. J., & Bárcena, M. (2020). A unifying structural and functional model of the coronavirus replication organelle: Tracking down RNA synthesis. *PLoS Biology*, 18(6), 1–25. <https://doi.org/10.1371/journal.pbio.3000715>
- Sola, I., Almazán, F., Zúñiga, S., & Enjuanes, L. (2015). Continuous and Discontinuous RNA Synthesis in Coronaviruses. *Annual Review of Virology*, 2, 265–288. <https://doi.org/10.1146/annurev-virology-100114-055218>
- Sola, I., Mateos-Gomez, P. A., Almazan, F., Zúñiga, S., & Enjuanes, L. (2011). RNA-RNA and RNA-protein interactions in coronavirus replication and transcription. *RNA Biology*, 8(2), 237–248. <https://doi.org/10.4161/rna.8.2.14991>
- Subissi, L., Imbert, I., Ferron, F., Collet, A., Coutard, B., Decroly, E., & Canard, B. (2014). SARS-CoV ORF1b-encoded nonstructural

- proteins 12-16: Replicative enzymes as antiviral targets. *Antiviral Research*, 101(1), 122–130. <https://doi.org/10.1016/j.antiviral.2013.11.006>
- Tagliamonte, M. S., Abid, N., Borocci, S., Sangiovanni, E., Ostrov, D. A., Kosakovsky Pond, S. L., Salemi, M., Chillemi, G., & Mavian, C. (2021). Multiple recombination events and strong purifying selection at the origin of SARS-CoV-2 spike glycoprotein increased correlated dynamic movements. *International Journal of Molecular Sciences*, 22(1), 1–16. <https://doi.org/10.3390/ijms22010080>
- Taylor, D. R., Puig, M., Darnell, M. E. R., Mihalik, K., & Feinstone, S. M. (2005). New Antiviral Pathway That Mediates Hepatitis C Virus Replicon Interferon Sensitivity through ADAR1. *Journal of Virology*, 79(10), 6291–6298. <https://doi.org/10.1128/jvi.79.10.6291-6298.2005>
- Wu, C., Liu, Y., Yang, Y., Zhang, P., Zhong, W., Wang, Y., Wang, Q., Xu, Y., Li, M., Li, X., Zheng, M., Chen, L., & Li, H. (2020). Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceutica Sinica B*, 10(5), 766–788. <https://doi.org/10.1016/j.apsb.2020.02.008>
- Zahn, R. C., Schelp, I., Utermöhlen, O., & von Laer, D. (2007). A-to-G Hypermutation in the Genome of Lymphocytic Choriomeningitis Virus. *Journal of Virology*, 81(2), 457–464. <https://doi.org/10.1128/jvi.00067-06>
- Zhao, S., & Chen, H. (2020). Modeling the epidemic dynamics and control of COVID-19 outbreak in China. *MedRxiv*, 1–9. <https://doi.org/10.1101/2020.02.27.20028639>
- Zhou, P., Yang, X. Lou, Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L., Chen, H. D., Chen, J., Luo, Y., Guo, H., Jiang, R. Di, Liu, M. Q., Chen, Y., Shen, X. R., Wang, X., ... Shi, Z. L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798), 270–273. <https://doi.org/10.1038/s41586-020-2012-7>
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., & Tan, W. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, 382(8), 727–733. <https://doi.org/10.1056/nejmoa2001017>