

Análisis de la mutación D614G en secuencia del genoma completo de SARS-CoV-2 en El Salvador

DOI 10.5377/alerta.v4i1.10683

Carlos Alexander Ortega Pérez¹, Noé Rigoberto Rivera^{2*}, Xochitl Sandoval López³, Carlos Enrique Hernández Ávila⁴.


1 y 2. Facultad de Medicina, Universidad de El Salvador, El Salvador.

3 y 4. Instituto Nacional de Salud, San Salvador, El Salvador.

*Correspondencia

✉ noe.rivera@ues.edu.sv

1.  0000-0003-1612-8386

2.  0000-0001-9553-0554

3.  0000-0002-0988-1313

4.  0000-0002-1230-0102



ACCESO ABIERTO

Analysis of the D614G mutation in whole genome sequence of SARS-CoV-2 in El Salvador

Citación recomendada:

Ortega CA, Rivera NR, Sandoval López X, Hernández CE.

Análisis de la mutación D614G en secuencia del genoma completo de SARS-CoV-2 en El Salvador. Alerta. 2021;4(1): DOI: 10.5377/alerta.v4i1.10683

Recibido:

21 de diciembre de 2020

Aceptado:

15 de enero de 2021

Publicado:

22 de enero de 2021

Contribución de autoría:

NRR¹, CAOP², XSL³, CEHA⁴.
Escritura del manuscrito.
NRR¹, CAOP², procesamiento de las muestras y análisis los datos en las plataformas bioinformáticas.

Conflicto de intereses:

Los autores declaran no tener conflictos de interés en el estudio ni haber recibido patrocinio de instituciones privadas; la fuente de financiamiento utilizada fueron recursos estatales designados para investigación.

Resumen

Introducción. El 18 de marzo se reporta el primer caso de infección por SARS-CoV-2 confirmado en El Salvador y durante el mes de octubre de 2020 se logra secuenciar el genoma de SARS-CoV-2 a partir de muestras obtenidas en el país. **Objetivo.** Analizar in silico las mutaciones detectadas en las secuencias aisladas en El Salvador. **Metodología.** Se utilizó la plataforma SOPHiA-DDM-V5.7.10, para la determinación de las variantes por mutaciones con sentido erróneo. Se utilizó la plataforma Nexclade beta v0.8.1; se visualizó y comparó la proteína S silvestre (D614: PDB ID: 6VXX) y de la variante mutada (D614G: PDB ID: 6XS6). Para el modelamiento y generación de imágenes de los detalles moleculares de las proteínas se utilizó Pymol-v1.7.2.3. **Resultados.** Los cristales de la proteína S silvestre y mutada muestra diferencias a nivel molecular, incluyendo la pérdida de interacciones entre el residuo G614 del dominio S1 y la treonina 859 de dominio S2, favoreciendo de esta manera la conformación abierta de la proteína S, la cual es necesaria para la interacción de S con el receptor ACE2. **Conclusión.** Los hallazgos confirman el predominio de la variante D614G en este grupo de secuencias, lo cual probablemente favorece su transmisibilidad, que puede explicarse por la configuración de los sitios de unión con receptor ACE2. El predominio mundial de la D614G y las evidencias de laboratorio y bioinformáticas publicadas hasta la fecha, apuntan hacia una posible mayor infectividad y transmisibilidad.

Palabras clave

SARS-CoV-2, D614G, NGS, 2019-nCoV, COVID-19.

Abstract

Introduction. The genome sequencing of indigenous samples of SARS-CoV-2 was carry out in October 2020. **Objective.** To analyze in silico the detected mutation in the isolated sequences in El Salvador. **Methodology.** The sequences were analyzed using the SOPHiA-DDM-V5.7.10 bioinformatic platform; the platform Nexclade beta v0.8.1 was used for the determination of variants due to missense mutation; the indigenous protein S (D614: PDB ID: 6VXX), was visualized and compared as well as the one from the mutant variant (D614G: PDB ID: 6XS6); Pymol-v1.7.2.3. was used in modeling and imaging generation of the proteins' molecular details. **Results.** The analysis of the wild protein S crystals and the mutant ones, show differences at molecular level, including the loss of interactions between the residue G614 of the S1 domain and the threonine 859 of the S2 domain, favoring in such manner, the open configuration of the protein S, which is necessary for the interaction between S with the ACE2 receptor. **Conclusion.** The global dominance of the D614G and the bioinformatic and laboratory evidences published up to date, show a possible higher infectivity and transmissibility conferred by the variant D614G detected in the sequence of SARS-CoV-2 in El Salvador.

Keywords

SARS-CoV-2, D614G, mutación, 2019-nCoV, COVID-19.

Introducción

El SARS-CoV-2 fue identificado por primera vez en Wuhan, Hubei, China en diciembre

del 2019 en un brote de un mercado de mariscos. El primer caso en EUA, fue confirmado en el estado de Washington D.C. el 20 de enero del 2020^{1,2}. El 18 de marzo

se reporta el primer caso de infección por SARS-CoV-2 confirmado en El Salvador³. A partir del primer caso de infección por SARS-CoV-2 en Wuhan, se analizó el genoma viral completo, revelando que el virus estaba más estrechamente relacionado con un grupo de coronavirus similares al SARS encontrado previamente en murciélagos en China⁴. Durante el año 2020 el SARS-CoV-2 se extendió de manera progresiva y sostenida por todo el mundo; en marzo se reportó el primer caso de infección por SARS-CoV-2 en El Salvador³ y en octubre de 2020 se logra secuenciar el genoma de SARS-CoV-2 a partir de muestras obtenidas de pacientes de El Salvador, con las secuencias obtenidas se analizó *in silico* las mutaciones detectadas en el país en particular la mutación D614G.

Genoma viral

Los coronavirus del género Betacoronavirus infectan solo a mamíferos pudiendo causar enfermedades respiratorias y gastroenteritis⁵. Los coronavirus HCoV-NL63, HCoV-229E, HCoV-OC43 y HKU1 clásicamente han causado afecciones respiratorias leves en humanos^{4,5}. Sin embargo, dos virus altamente patógenos, SARS-CoV y MERS-CoV causan en humanos síndrome respiratorio severo agudo en humanos^{5,6}.

Los coronavirus forman partículas esféricas o pleomórficas envueltas de 100 a 160 nm de diámetro, contienen un genoma de ARN monocatenario (ssRNA) de sentido positivo, con un tamaño entre 27 a 32 kb. El genoma de SARS-CoV-2, de aproximadamente 30 kb, codifica para 14 marcos de lectura abierto (ORF del inglés Open Reading Frame). En el extremo 5' del genoma viral, los primeros son los ORF1a y ORF1ab, que ocupan aproximadamente dos tercios de la longitud total del genoma, codifican para poliproteínas que ayudan a controlar los ribosomas del huésped para su propio proceso de traducción, los que se procesan autoproteolíticamente en 16 proteínas no estructurales que forman el complejo replicasa-transcriptasa. Un desplazamiento del marco de lectura entre ORF1a y ORF1b conduce a la producción de dos polipéptidos: pp1a y pp1ab. Estos polipéptidos son procesados por proteasas similares a quimotripsina codificada viralmente (3CL pro) o proteasa principal (M pro) y una o dos proteasas similares a la papaína en 16 NSPs (NSP1-NSP16). Este complejo replicasa-transcriptasa consta de múltiples enzimas, incluida la proteasa similar a la papaína (NSP3), la proteasa principal (NSP5), el complejo primasa

NSP7-NSP8, la ARN polimerasa primaria dependiente de ARN (NSP12), una helicasa-trifosfatasa (NSP13), una exoribonucleasa (NSP14), una endonucleasa (NSP15), N7- y 2'O- metiltransferasa (NSP10 y NSP16). En el extremo 3' otros ORFs en el tercio distal del genoma codifican para al menos cuatro proteínas estructurales principales: proteína de espícula (S), glicoproteína de membrana (M), envoltura (E) y nucleocápside (N), además nueve factores accesorios putativos⁷⁻¹⁰.

Se sabe que el SARS-CoV-2 tiene al menos seis proteínas accesorias (3a, 6, 7a, 7b, 8 y 10) según la anotación actual (GenBank: NC_045512.2). Sin embargo, aún no se ha verificado experimentalmente la expresión de los ORFs y no está claro qué genes accesorios se expresan realmente a partir de este genoma compacto¹¹. Muchas de estas proteínas funcionales y estructurales virales tienen una función ya definida; sin embargo, algunas aún están pendientes de definir su función¹⁰.

Estructura y función de la glucoproteína espícula (S)

La entrada del SARS-CoV-2 en la célula hospedadora es mediada por la glucoproteína transmembrana Spike (S) que forma homotrímeros (Figura 1) que protruyen en la superficie viral¹².

Con un tamaño de 180-200 kDa, la proteína S consta de un extremo N-terminal externo, un dominio transmembrana (TM) anclado en la membrana viral y un extremo corto C-terminal (CT)¹³. El número total de residuos de aminoácidos de la proteína S es de 1273 aa y consta de un péptido señal (aminoácidos 1-13) ubicado en el extremo N-terminal, la subunidad S1 (residuos 14-685) y la subunidad S2 (residuos 686-1273). En la subunidad S1 se encuentra un dominio N-terminal (residuos 14-305) y una región de unión al receptor de dominio (RBD, por sus siglas en inglés Receptor Binding Domain) (319-541 residuos). En la subunidad S2 encontramos el péptido de fusión (residuos 788-806), la secuencia de repetición de heptapéptido 1 (HR1) (residuos 912-984) HR2 (residuos 1163-1213), el dominio transmembrana o TM (residuos 1213-1237), el dominio citoplasmático (residuos 1237-1273) y el péptido de fusión (FP) (residuos 788-806)¹⁴. Los trímeros de la proteína S forman visualmente los característicos bulbos en forma de corona que rodea la partícula viral¹⁵. La estructura atómica tridimensional del homotrímero S del SARS-CoV-2 (figura 1.A) ha sido resuelta por microscopía crioelectrónica, revelando diferentes conformaciones del dominio RBD de la proteína S en dos estados: abierto y cerrado^{16,17}.

El RBD de la región distal de S1 es un dominio de unión al receptor celular y contribuye además a la estabilización del estado de prefusión de la subunidad S2, anclada a la membrana que contiene la maquinaria de fusión^{18,19}. Para todos los CoVs, S se escinde adicionalmente por proteasas del hospedador en el denominado sitio S20, ubicado antes del péptido de fusión^{20,21}. Se ha propuesto que esta escisión activa la proteína para que se fusione con la membrana a través de amplios e irreversibles cambios conformacionales²². Como resultado, la entrada del coronavirus en las células susceptibles es un proceso complejo que requiere la acción concertada de unión al receptor y procesamiento proteolítico de la proteína S para lograr finalmente la fusión virus-célula.

Cada una de los RBD de los monómeros o protómeros de proteína S1 que forman la espícula del SARS-CoV-2 puede individualmente adoptar una conformación abierta o cerrada^{12,16}.

El sitio de unión al receptor en la proteína S está parcialmente oculto en la conformación cerrada; en la forma abierta, por el contrario, dicho sitio de interacción queda expuesto y puede interactuar con el receptor ACE2, permitiendo la unión del virus a la membrana de la célula y facilitando finalmente la entrada del virus^{23,24}. En la proteína S silvestre el porcentaje de protómeros que se encuentran en configuración abierta es del 18 % y cerrada 82 %, lo anterior favorecido por las interacciones que existen entre el aspartato 614 de S1 (D614) y la treonina 859 (T859) de S2, de esta manera el residuo D614 funciona como un interruptor a distancia (efecto alostérico) que mantiene mayoritariamente cerrada la proteína S limitando la interacción del RBD con el receptor ACE2²⁵. (Figura 1.B)

Metodología

Las secuencias analizadas registradas en la plataforma Global Initiative on Sharing All Influenza Data (GISAID por sus siglas en inglés)²⁶ fueron obtenidas a partir de eluidos de ARN de secreción nasofaríngea de pacientes positivos a la prueba de qRT-PCR para SARS-CoV-2, diagnosticados en el Laboratorio Nacional de Salud Pública del Instituto Nacional de Salud (INS) y la Sección de Virología y Microbiología Molecular, Departamento de Microbiología de la Facultad de Medicina, Universidad de El Salvador.

La secuenciación fue realizada en la Sección de Virología y Microbiología Molecular, Departamento de Microbiología de la Facultad de Medicina, Universidad de El Salvador, empleando el método de secuenciación con enfoque de amplicones. Se utilizó una

variante de reacciones de PCR multiplex de dos grupos con el kit de Paragon CleanPlex® SARS-CoV-2 by SOPHiA. Para la secuenciación se utilizó la plataforma MiniSeq Illumina según consta en los registros de GISAID para las secuencias de El Salvador.

Para la determinación de las variantes por mutaciones con sentido erróneo se utilizó la plataforma bioinformática SOPHiA-DDM-V5.7.10. Para la verificación de las mutaciones, las secuencias fueron analizadas en la plataforma GISAID²⁷ y Nexclade beta v0.8.1.

Los SNPs con sentido erróneo se tabularon asociados al gen afectado y el respectivo cambio de aminoácido. Obtenidos los SNPs se escogió la variante D614G encontrada en todas las secuencias que actualmente está siendo ampliamente estudiadas, Para ello se analizó mediante la visualización y comparación de la proteína S silvestre (D614 : PDB ID: 6VXX) y de la variante mutada (D614G : PDB ID: 6XS6), que fueron descargadas del Protein Data Bank²⁸ para el modelamiento y generación de imágenes de los detalles moleculares de las proteínas silvestre y mutada, se utilizó Pymol-v1.7.2.3.

Consideraciones éticas

El protocolo de investigación fue sometido y aprobado por el comité de ética local del Instituto Nacional de Salud. Las muestras fueron manejadas con un número de referencia para resguardar la identidad de las personas de las que procedían las muestras. Los datos de la secuenciación están disponibles en la plataforma GISAID bajo los siguientes registros: EPI_ISL_672570, EPI_ISL_672572, EPI_ISL_672573 ,EPI_ISL_671974, EPI_ISL_671978, EPI_ISL_672012.

Resultados

Mutaciones con sentido erróneo en el genoma de SARS-CoV-2 aislados en El Salvador

Se analizaron los genes E, M, N, ORF10, ORF1a, ORF1ab, ORF3a, ORF6, ORF7a, ORF7b, ORF8, S, dentro de los cuales se identificaron un total de 27 polimorfismos de un solo nucleótido (SNP). Desde 10 y hasta 15 mutaciones con sentido erróneo en cada una de las seis secuencias, algunos de los SNPs comunes entre las secuencias. Además, la mayoría de los SNP identificados fueron sin sentido (missense) implicando el cambio de aminoácido. Las secuencias, basadas en los SNP identificados, se agruparon en un solo patrón variante, impulsada principalmente por SNP ubicados en la región S, NSP2, NS3 y gen S, presentando las muta-

ciones A23403G (S^{D614G}), C1059T ($NSP2^{T265I}$) y G25563T ($NS3^{Q57H}$). Se puede observar que la mutación D614G está presente en todas las secuencias. (Ver tabla 1)

Discusión

Las seis secuencias aisladas de pacientes salvadoreños (Tabla 1) fueron clasificadas en el clado Nextstrain 20C, presentando las mutaciones A23403G (S^{D614G}), C1059T ($NSP2^{T265I}$) y G25563T ($NS3^{Q57H}$). Las variantes del SARS-CoV-2 con la mutación S^{D614G} (clados GISAID G, GH y GR) en la proteína S predominan actualmente a nivel mundial²⁹.

Papel de la mutación D614G en la proteína S

La mutación S^{D614G} tiene dos consecuencias. Primero, el residuo G614 dificulta la formación de interacciones (puentes de hidrógeno) con la treonina 859 (T859) en S2, debilitando de esa manera la unión entre dos

protómeros diferentes y reduciendo la compactación del trímero de proteína S.

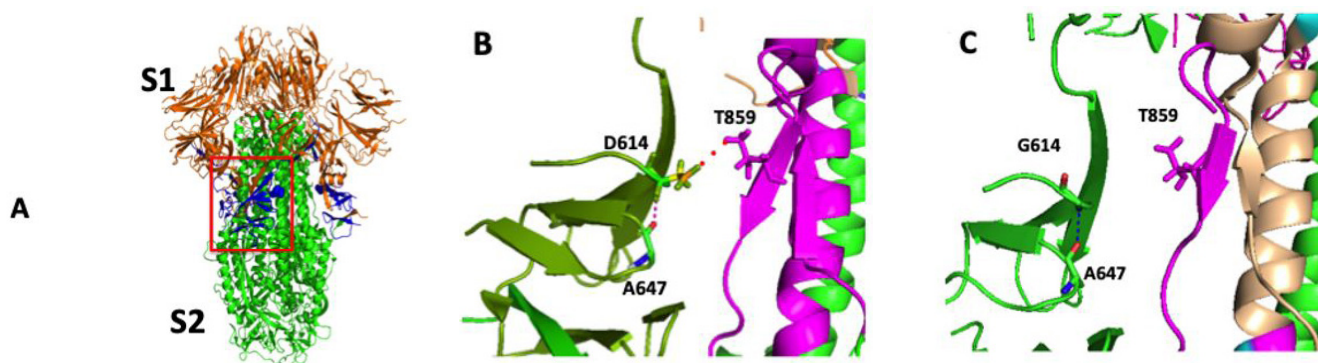
Segundo, provoca la reducción de la distancia entre el residuo G614 y A647 dentro de cada protómero S, favoreciendo el estado abierto del dominio RBD²⁵ (Figura 1.C). A diferencia de la variante S^{D614} cuyo porcentaje de protómeros en conformación abierta es del 18 %¹⁷, en la S^{G614} el 54 % de los protómeros se encuentran en conformación abierta encontrándose el dominio RBD más accesible para interactuar con el receptor ACE2²⁵.

Los estudios con modelos celulares que expresan ACE2 mostraron que los viriones que llevan S^{G614} entran en células que expresan ACE2 más eficientemente que aquellos con S^{D614} . Este aumento de la entrada se correlaciona con menos desprendimiento del dominio S1 y una mayor incorporación de proteína S en el virión. Por otro lado, las variantes no están relacionadas con el desarrollo de un cuadro clínico más severo³⁰, ni afectan la potencia de anticuerpos contra SARS-CoV-2³¹.

Tabla 1. Mutaciones con sentido erróneo encontradas

Identificación GISAID	EPI_ISL_672570	EPI_ISL_672572	EPI_ISL_672573	EPI_ISL_671974	EPI_ISL_671978	EPI_ISL_672012
S	D614G	D614G	D614G	A623S, D614G	D614G	D614G, S704L
N	R209K	R209K	R209K	R209K	R209K	P67S, P199L
NS3	P258S, Q57H	P258S, Q57H	P258S, Q57H	P258S, Q57H	P258S, Q57H	G172V, Q57H
NS7a	T39I	T39I	T39I	T39I	T39I	-
NS8	S24L, V62L	S24L	S24L	S24L	S24L	S24L
NSP2	T85I, T265I	T85I, T265I	S430L, T85I, T265I	T85I, T265I	T85I, T265I	T85I, T265I
NSP3	-	-	K384N	-	-	H727Y
NSP5	-	-	-	-	-	L89F
NSP9	-	-	S13F	-	-	-
NSP12	K59N, P323L	K59N, P323L	K59N, P323L	K59N, P323L	K59N, P323L	P323L
NSP13	-	-	-	L256F, N124S	L256F, N124S	-
NSP14	-	-	-	-	-	N129D
NSP15	-	-	-	-	-	V320L
NSP16	-	-	-	-	-	R216C

Figura 1. Estructura tridimensional de la proteína S de SARS-CoV-2.



Este estudio abarca únicamente seis secuencias. Es necesario continuar la secuenciación de muestras representativas del país que permitan establecer la prevalencia de la misma y explicar la transmisibilidad, además del comportamiento de las variedades clínicas de la enfermedad que se han presentado en el país.

Conclusiones

En El Salvador, el reporte de las primeras seis secuencias completas ha permitido clasificar los genomas en el clado 20C según Nextstrain y el clado GH según GISAID, ubicándolos en los linajes pangolín B.1.2. y B.1.370, que son predominantes en EE.UU.

Las mutaciones detectadas en las secuencias aisladas en El Salvador afectan principalmente a los genes de las proteínas: S, N, NS3, NS7a, NS8, NSP2 y NSP12.

Al igual que a nivel mundial, los hallazgos confirman también el predominio de la variante D614G en este grupo de secuencias, que puede explicarse por la configuración de los sitios de unión con receptor ACE2, lo cual probablemente favorece su transmisibilidad.

Es necesario continuar con los estudios de secuenciación de genomas completos para confirmar de manera definitiva la prevalencia de la variante D614G en El Salvador y otras que puedan presentarse. Por otro lado, también es necesario continuar atentos a las evidencias clínicas, de laboratorio y bioinformáticas que confirmen el papel de la mutación D614G en la infectividad y transmisibilidad de SARS-CoV-2. El análisis de las otras mutaciones encontradas será objeto de nuevas publicaciones que son necesarias para comprender el comportamiento epidemiológico y clínico de la enfermedad en el país, con el objeto de emprender estrategias para su abordaje integral, contribuir a la vigilancia genómica a nivel mundial y a generar información útil para el diseño de vacunas y fármacos blanco.

Conflicto de intereses

Los autores declaran no tener conflictos de interés en el estudio ni haber recibido patrocinio de instituciones privadas. La fuente de financiamiento utilizada fue a partir de recursos estatales designados para investigación.

Agradecimientos

Se reconoce el apoyo de las máximas autoridades de la Universidad de El Salvador, Instituto Nacional de Salud y Ministerio de Salud.

Agradecimientos especiales a la Sección de Virología del Departamento de Microbiología, Facultad de Medicina, Universidad de El Salvador, y a la Sección de Virología del Laboratorio Nacional de Salud Pública de El Salvador.

Referencias Bibliográficas

1. Harcourt J, Tamin A, Lu X, Kamili S, Sakthivel SK, Murray J et al. Severe acute respiratory syndrome coronavirus 2 from patient with coronavirus disease, United States. *Emerging Infectious Diseases*. 2020;26(6):1266-1273. DOI:10.3201/EID2606.200516
2. Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, et al. First Case of 2019 Novel Coronavirus in the United States. *N. Engl. J. Med*. 2020;382(10):929-936. doi:10.1056/nejmoa2001191
3. Gobierno de El Salvador. COVID-19 Reporte Diario. Gobierno de la Republica de El Salvador. 2021. Fecha de consulta: 21 de enero de 2021. Disponible en: <https://covid19.gob.sv/diarios/>
4. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579(7798):265-269. doi:10.1038/s41586-020-2008-3
5. Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, et al. Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends Microbiol*. 2016;24(6):490-502. doi:10.1016/j.tim.2016.03.003
6. Woo PCY, Lau SKP, Lam CSF, Lau CCY, Tsang AKL, Lau JHN, et al. Discovery of Seven Novel Mammalian and Avian Coronaviruses in the Genus Deltacoronavirus Supports Bat Coronaviruses as the Gene Source of Alphacoronavirus and Betacoronavirus and Avian Coronaviruses as the Gene Source of Gammacoronavirus and Deltacoronavi. *J. Virol*. 2012;86(7):3995-4008. doi:10.1128/jvi.06540-11
7. Chen Y, Liu Q, Guo D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J. Med. Virol*. 2020;92(4):418-423. doi:10.1002/jmv.25681
8. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol*. 2019;17(3):181-192. doi:10.1038/s41579-018-0118-9
9. Fung SY, Yuen KS, Ye ZW, Chan CP, Jin DY. A tug-of-war between severe acute respiratory syndrome coronavirus 2 and host antiviral defence: lessons from other pathogenic viruses. *Emerg. Microbes Infect.* 2020;9(1):558-570. doi:10.1080/22221751.2020.1736644
10. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, O'Meara MJ, et al. A SARS-

- CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug- Repurposing. *BioRxiv Prepr. Serv. Biol.* 2020;(2020.03.22.002386.). doi:<https://doi.org/10.1101/2020.03.22.002386>
11. Kim D, Lee JY, Yang JS, Kim JW, Kim VN, Chang H. The Architecture of SARS-CoV-2 Transcriptome. *Cell.* 2020;181(4):914-921. e10. doi:10.1016/j.cell.2020.04.011
 12. Park YJ, Walls AC, Wang Z, Sauer MM, Li W, Tortorici MA, et al. Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors. *Nat. Struct. Mol. Biol.* 2019;26(12):1151–1157. doi:10.1038/s41594-019-0334-7
 13. Bosch BJ, van der Zee R, de Haan CAM, Rottier PJM. The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. *J. Virol.* 2003;77(16):8801–8811. doi:10.1128/jvi.77.16.8801-8811.2003
 14. Xia S, Zhu Y, Liu M, Lan Q, Xu W, Wu Y, et al. Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. *Cell. Mol. Immunol.* 2020;17(7):765–767. doi:10.1038/s41423-020-0374-2
 15. Tang T, Bidon M, Jaimes JA, Whittaker GR, Daniel S. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. *Antiviral Research.* 2020;178:104792. doi:10.1016/j.antiviral.2020.104792
 16. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science.* 2020;367(6483):1260–1263. doi:10.1126/science.aax0902
 17. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Velesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell.* 2020;181(2):281-292.e6. doi:10.1016/j.cell.2020.02.058
 18. Gui M, Song W, Zhou H, Xu J, Chen S, Xiang Y, et al. Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res.* 2017;27(1):119–129. doi:10.1038/cr.2016.152
 19. Kirchdoerfer RN, Cottrell CA, Wang N, Pallesen J, Yassine HM, Turner HL, et al. Pre-fusion structure of a human coronavirus spike protein. *Nature.* 2016;531(7592):118–121. doi:10.1038/nature17200
 20. Madu IG, Roth SL, Belouzard S, Whittaker GR. Characterization of a Highly Conserved Domain within the Severe Acute Respiratory Syndrome Coronavirus Spike Protein S2 Domain with Characteristics of a Viral Fusion Peptide. *J. Virol.* 2009;83(15):7411–7421. doi:10.1128/jvi.00079-09
 21. Millet JK, Whittaker GR. Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus Research.* 2015;202:120–134. doi:10.1016/j.virusres.2014.11.021
 22. Park JE, Li K, Barlan A, Fehr AR, Perlman S, McCray PB, et al. Proteolytic processing of middle east respiratory syndrome coronavirus spikes expands virus tropism. *Proc. Natl. Acad. Sci. U. S. A.* 2016;113(43):12262–12267. doi:10.1073/pnas.1608147113
 23. Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, et al. Structural basis of receptor recognition by SARS-CoV-2. *Nature.* 2020;581(7807):221–224. doi:10.1038/s41586-020-2179-y
 24. Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science.* 2020;367(6485):1444–1448. doi:10.1126/science.abb2762
 25. Yurkovetskiy L, Wang X, Pascal KE, Tomkins-Tinch C, Nyalile TP, Wang Y, et al. Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell.* 2020;183(3):739-751.e8. doi:10.1016/j.cell.2020.09.032
 26. GISAID - Initiative. Accessed August 23, 2020. <https://www.gisaid.org/>
 27. GISAID - phylodynamics. Accessed December 22, 2020. Disponible en: www.gisaid.org/epiflu-applications/phylodynamics/
 28. Bank RPD. RCSB PDB: Homepage. Accessed December 22, 2020. <https://www.rcsb.org/>
 29. Isabel S, Graña-Miraglia L, Gutierrez JM, Bundalovic-Torma C, Groves HE, Isabel MR, et al. Evolutionary and structural analyses of SARS-CoV-2 D614G spike protein mutation now documented worldwide. *Sci. Rep.* 2020;10(1):1–20. doi:10.1038/s41598-020-70827-z
 30. Zhang L, Wang S, Ren Q, Yang J, Lu Y, Zhang L, et al. Genome-wide variations of SARS-CoV-2 infer evolution. *MedRxiv Prepr.* 2020:1–17. doi: <https://doi.org/10.1101/2020.04.27.20081349>
 31. Zhang L, Jackson CB, Mou H, Ojha A, Peng H, Quinlan BD, et al. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat. Commun.* (2020):1–9. doi:10.1038/s41467-020-19808-4