



MINISTERIO
DE CIENCIA
E INNOVACIÓN



Instituto de Salud Carlos III

INFORME DEL GRUPO DE ANÁLISIS CIENTÍFICO DE CORONAVIRUS DEL ISCI (GACC-ISCI)

EVOLUCIÓN DEL CORONAVIRUS SARS-CoV-2

** Este informe está realizado con la evidencia científica disponible en la fecha de su elaboración y podrá ser actualizado si surgen nuevas evidencias*

RESUMEN DIVULGATIVO

El [genoma de un organismo](#) es el conjunto de toda su información genética, el 'libro' que define sus principales características biológicas. El primer genoma del coronavirus SARS-CoV-2 se obtuvo el pasado mes de enero y fue el primer paso para comprender mejor cómo se comporta y actúa el virus. Desde entonces [se han secuenciado](#) más de 40.000 genomas del SARS-CoV-2 [en todo el mundo](#), [una información](#) que está permitiendo [rastrear cómo se propaga el virus](#) gracias a estudios de genética y epidemiología molecular.

Todos los virus van generando copias de su genoma mientras infectan a otros organismos. En este proceso se van produciendo pequeños cambios, mutaciones genéticas en el genoma cuyo análisis permite trazar cómo se transmite el virus entre personas. Al investigar estas mutaciones en el SARS-CoV-2, los científicos han podido establecer lo que se conoce como 'clusters' filogenéticos del coronavirus, diferentes tipos o 'ramas' del virus que explican su origen, evolución y difusión.

De esta manera, ya hay bastante información sobre cómo se ha propagado el virus por todo el mundo, y sobre qué mutaciones y características tiene en diferentes localizaciones geográficas. Hasta el momento se han diferenciado varias 'familias' del nuevo coronavirus, denominadas clados filogenéticos, caracterizadas por diferentes mutaciones. Todos los grandes clados del virus, que ayudan a explicar su origen y distribución, se han encontrado en prácticamente todos los países del mundo; todos los clados están en casi todos los países, con variaciones en la frecuencia de cada uno.

Una de las variantes del SARS-CoV-2 se ha convertido en la forma genética mayoritaria en muchos países del mundo, y en Europa en particular. [El análisis de las mutaciones que va sufriendo el virus](#) también está permitiendo investigar si, según pasa el tiempo, su capacidad de transmisión e infección se atenúa o se hace más fuerte. Por el momento hay escasas evidencias de que algunas de las variantes que se conocen del SARS-CoV-2 puedan ser más o menos agresivas o virulentas.

Las características genéticas y evolución del virus se siguen estudiando. El análisis de cómo se transmite, mediante una disciplina conocida como epidemiología genómica, es fundamental para conocer la diversidad del virus en un territorio concreto, evaluar su propagación y facilitar la toma de decisiones y medidas de contención para evitar su expansión.

INFORME COMPLETO

El coronavirus SARS-CoV-2, descubierto en enero de 2020 tras aislarse de muestras de pacientes afectados por una nueva enfermedad ahora conocida como COVID-19, evoluciona y sufre cambios genéticos, como todos los virus. Conocer estos cambios, que explican su comportamiento, es fundamental para mejorar el manejo del virus y el abordaje de la enfermedad. Desde el hallazgo del SARS-CoV-2 hasta ahora, [la secuenciación de su genoma](#) y el conocimiento de las diferentes variantes que circulan por el mundo está permitiendo conocer más sobre su origen, influencia y distribución.

Primeros casos de la infección por SARS-CoV-2 en China

Un trabajador de un mercado de marisco de la ciudad china de Wuhan comenzó a encontrarse enfermo y fue ingresado en el hospital de su ciudad el 26 de Diciembre del 2019 con neumonía severa e insuficiencia respiratoria aguda [1]. Tras descartar la presencia de todos los patógenos respiratorios habituales, se tomó una muestra del paciente para secuenciar el posible agente etiológico. Apenas tres semanas después, se consiguió obtener el genoma del que pasaría a ser el séptimo representante de la familia coronavirus con capacidad de infectar a los humanos.

Este virus acabó denominándose SARS-CoV-2. La secuenciación de los 29,899 nucleótidos de su genoma supuso una referencia sobre la que los investigadores pudieran empezar a trabajar [2]. Así comenzó una carrera contrarreloj en la que hasta el momento se han conseguido secuenciar más de 56.000 genomas completos de todo el mundo. Más de la mitad de las secuencias (36.404) pertenecen a virus aislados de personas infectadas en Europa y 1674 en España [3].

El hecho de tener la información genómica de estos virus, junto con información epidemiológica básica como el dónde y cuándo se ha obtenido la muestra, permite a los investigadores rastrear cómo se propaga el virus por el mundo gracias a la epidemiología molecular/genómica. Esto es posible por una característica de los virus relacionada con su necesidad de utilizar la maquinaria celular del organismo al que infectan para poder multiplicarse y sobrevivir.

Sin embargo, esas copias que se van generando no siempre son copias perfectas sino que a veces se comenten errores y poco a poco se van modificando alguno de los 29.899 nucleótidos que conforman su genoma. Estos pequeños errores o mutaciones que se van produciendo en el genoma son las que permiten trazar cómo se transmite el virus entre personas.

En el caso de los primeros casos de China, los genomas de los virus obtenidos de otros trabajadores del mercado que también habían enfermado fueron secuenciados. Estos genomas resultaron ser copias exactas del genoma encontrado en el primer paciente, salvo por un pequeño número de nucleótidos, entre 1 y 5. Esta similitud entre las secuencias permite definir que esos virus proceden de un mismo ancestro y forman lo que se conoce como un cluster filogenético.

Una vez que se han secuenciado un número suficiente de genomas, se puede estimar cuánto es capaz de mutar el virus, es decir, cuantas mutaciones acumula el virus en un tiempo

determinado (tasa de evolución). En concreto, se ha estimado que las más de 100 primeras secuencias disponibles a principios de febrero del 2020 ya eran suficientes para calcular de manera fiable dicha tasa de evolución, ya que ésta no se veía alterada al añadir nuevas secuencias al análisis [4].

Se pudo estimar que SARS-CoV-2 acumulaba mutaciones a un ritmo de entre 1.19 y 1.31×10^{-3} sustituciones/sitio/año, parecido al de los otros coronavirus epidémicos, ya que en el caso de SARS-CoV se estima una tasa entre 0.80×10^{-3} y 2.38×10^{-3} y para MERS-CoV entre 0.88×10^{-3} y 1.37×10^{-3} [5, 6, 7]. Conocer este ritmo de evolución, que supondría aproximadamente la acumulación de una mutación cada 10 días, permite estimar la localización geográfica y temporal para el ancestro de un cluster filogenético concreto. Así, analizando los primeros genomas secuenciados, se ha podido determinar que el origen de la epidemia tuvo lugar a finales de noviembre del año 2019, [4, 7-10].

Primeros casos de la infección por SARS-CoV-2 en Europa y España

Con el paso del tiempo han ido apareciendo diferentes mutaciones a lo largo del genoma del virus, lo que da valiosas pistas para rastrear la propagación del SARS-CoV-2 por todo el mundo. Uno de los primeros casos de infección por SARS-CoV-2 detectado en Europa tuvo lugar cerca de Múnich, Alemania. Fue el caso de una mujer que aterrizó el 22 de enero en el aeropuerto de Múnich y que había estado en contacto con sus padres residentes en Wuhan. Esta mujer era trabajadora de una empresa alemana en la que pocos días después enfermaron varios trabajadores. El genoma del virus detectado en uno de estos empleados era una copia casi idéntica de otros virus secuenciados en Shanghai y muy parecida al primer genoma secuenciado del mercado de Wuhan [11].

En concreto, el genoma del virus secuenciado en Alemania (BavPat1) presentaba las mutaciones C3037T, en el gen ORF1ab y A23403G, que provocaba el cambio de aminoácidos D614G en el gen de la espícula del virus, con respecto al genoma de referencia (Wuhan-Hu-1). Este cambio de aminoácidos ya había sido detectado en otros virus encontrados en Shanghai y acabaría siendo una de las mutaciones encontradas en la mayoría de los genomas de los virus secuenciados en Europa durante las siguientes semanas [11]. La investigación epidemiológica de esta cadena de transmisión permitió detectar que uno de los 12 casos relacionados con este primer positivo de Europa se acabaría convirtiendo en el primer caso detectado en España, en un individuo que voló de Múnich a Las Islas Canarias el día 28 de enero del 2020 [11].

Otro de los primeros casos de infección por SARS-CoV-2 detectados en Europa tuvo su origen en una conferencia de una compañía de gas en Singapur durante el 20-22 de enero del 2020 a la que acudieron más de 100 personas de todo el mundo. El 9 de febrero, por lo menos 7 asistentes habían dado positivo a la infección por SARS-CoV-2 en Singapur, Malasia, Corea del Sur y Reino Unido. Uno de ellos está relacionado con un brote europeo de al menos otras 13 personas con diagnóstico positivo a COVID-19 y que tuvo su origen en una estación de esquí francesa, afectando a personas de Reino Unido, Francia y España [12]. En concreto, este caso está relacionado con el segundo positivo detectado en España en la isla de Mallorca [12].

El Reino Unido es el país que mayor esfuerzo en secuenciación de genomas completos ha realizado a nivel mundial, con más de 20.000 secuencias [13]. Esta información, junto con los

genomas de virus procedentes de infecciones en todo el mundo, ha permitido a los investigadores realizar una primera aproximación sobre los eventos de importación y establecimiento de linajes del virus en todo el país. Han podido detectar al menos 1.356 introducciones independientes del virus en el Reino Unido desde otras regiones del mundo (aunque alertan de una posible infraestimación del número). En torno al 25% de esas introducciones parecen haberse extinguido ya que no se han encontrado nuevas secuencias que pertenezcan a esos linajes durante un periodo de 4 semanas [13]. El 80% de las importaciones desde otros países se produjeron principalmente durante el mes de marzo del 2020, alcanzando su pico a mediados del mismo mes.

Calculando el tiempo de los ancestros comunes más recientes de todos esos linajes determinaron como mediana el 25 de marzo de 2020 (rango intercuartil 17 Marzo – 1 Abril, 2020) [13]. Además, han determinado que una cada tres importaciones en el Reino Unido proceden de viajeros procedentes de España desde el 16 de febrero, pero principalmente durante la primera quincena de marzo, fecha a partir de la cual comenzaron a descender las importaciones desde España hasta llegar a desaparecer la primera semana de abril del 2020 [13]. Las 50 cadenas de transmisión más antiguas detectadas en este trabajo y 49 de las 50 cadenas con mayor número de secuencias están relacionadas con variantes G614, es decir, esta variante llegó antes y con mayor intensidad a Reino Unido, lo que podría sugerir un efecto fundador en un país donde dicha variante es claramente mayoritaria.

Otro ejemplo es el caso específico de Escocia, donde se ha conseguido secuenciar el genoma del 20% (n=452) de todos los individuos con diagnóstico de COVID-19 disponibles a día 1 abril en toda el país (n=2310) [14]. Esta información ha permitido obtener una buena fotografía sobre la epidemia y las importaciones y dispersión del virus en Escocia. De este modo, se han detectado 113 introducciones independientes. El 51% de los clusters identificados se asociaron a introducciones no detectadas y relacionadas con virus que circulaban ya en otros países europeos como España, Italia y Austria. Los autores defienden que introducciones no documentadas ocurrieron antes de los primeros casos detectados [14].

Epidemiología genómica del SARS-CoV-2: diferentes nomenclaturas

Gracias a un esfuerzo sin precedentes se han conseguido secuenciar más de 40.000 genomas completos del SARS-CoV-2 en países de todo el mundo. Esto ha permitido que plataformas como [Nextstrain](#) hayan podido mostrar prácticamente a tiempo real la epidemiología genómica del virus y como ha ido transmitiéndose a lo largo del tiempo por los diferentes países afectados (Figura 1) [15].

Actualmente, *Nextstrain* define 5 grandes clados filogenéticos para clasificar los genomas que se van secuenciando y que se nombran en función del año estimado en el que emergieron (19 o 20) seguido de una letra:

- 19A: que se considera el clado raíz del que surgen todos los demás y que alcanzó una frecuencia global entre el 47%-65% en enero del 2020.
- 19B: caracterizado por las mutaciones C8782T y T28144C y que también alcanzó una alta prevalencia en Asia en enero del 2020 (28-33%).



GOBIERNO
DE ESPAÑA

MINISTERIO
DE CIENCIA
E INNOVACIÓN



Instituto de Salud Carlos III

- 20A: caracterizado por las mutaciones C14408T y A23403G, alcanzando una frecuencia global del 41-46% en Abril-Mayo del 2020 principalmente por países de Norteamérica, Europa y Asia.
- 20B: con las mutaciones consecutivas G28881A, G28882A y G28883C, que alcanzó una prevalencia en torno al 20% en Marzo-Abril principalmente por secuencias de Europa.
- 20C: caracterizado por C1059T y G25563T, alcanzando una frecuencia global en torno al 20% en Abril principalmente por secuencias de EEUU.

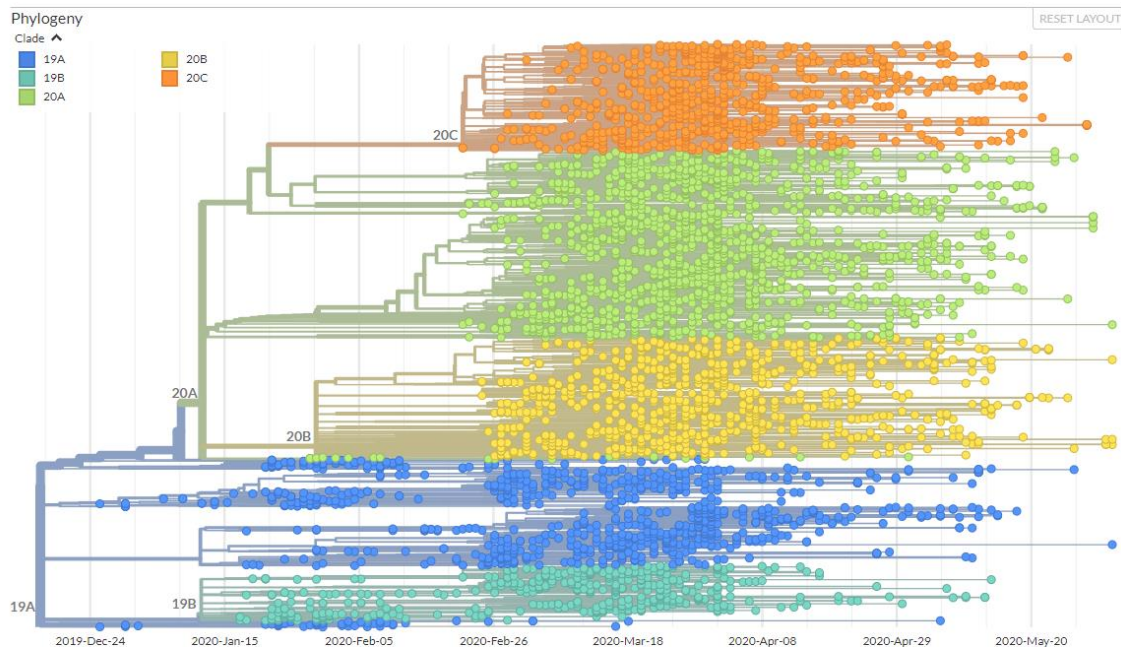


Figura 1. Principales clados filogenéticos definidos por Nextstrain

Por su parte, [la base de datos de GISAID](#), en la que se depositan gran parte de los genomas que se van secuenciando a nivel global, clasifica dichos genomas de una manera diferente [16]. Los grandes clados definidos por GISAID son:

- Clado S: caracterizado por la presencia de la mutación L84S en la proteína NS8 del virus.
- Clado V: caracterizado por la presencia de la mutación G251V en la proteína NS3 del virus.
- Clado G: caracterizado por la presencia de la mutación D614G en la espícula del virus.
- Clado GH: que también tiene la mutación D614 en la espícula del virus pero además tiene la mutación Q57H en la proteína NS3.
- Clado GR: que también tiene la mutación D614 en la espícula del virus pero además tiene la mutación G204R en la proteína de la nucleocápside del virus.



GOBIERNO
DE ESPAÑA

MINISTERIO
DE CIENCIA
E INNOVACIÓN



Instituto de Salud Carlos III

Además, existe una tercera nomenclatura propuesta por un importante grupo de virólogos evolutivos basada en la definición de linajes jerarquizados A y B que propone un sistema dinámico para definir la aparición de nuevos brotes locales con importancia epidemiológica [17]. Aunque son sistemas diferentes, existen ciertas equivalencias (no perfectas) entre los diferentes clados definidos por los tres sistemas de nomenclatura. De este modo, el linaje B incluiría virus de los clados G, GH y GR del sistema de GISAID y los clados 20A, 20B y 20C del sistema de *Nextstrain*. Por su parte, el linaje A incluiría los clados S y V del sistema de GISAID y los clados 19A y 19B de *Nextstrain*.

Existe una iniciativa pública española denominada [NextSpain](#) que facilita el análisis e interpretación de resultados sobre epidemiología y evolución del SARS-CoV-2 en España siguiendo los procedimientos de *Nextstrain*. En la Figura 2 puede verse como se distribuyen los genomas españoles en los clados definidos por *Nextstrain*.

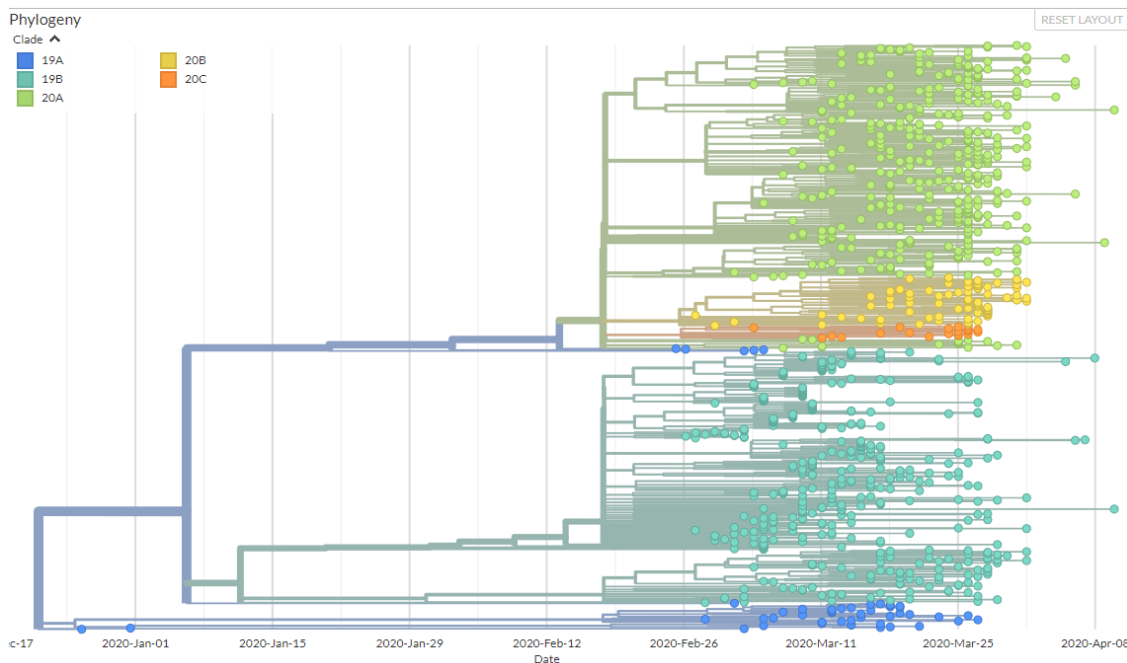


Figura 2. Principales clados filogenéticos de secuencias de España definidos por *Nextstrain* y alojados en *NextSpain*.

Situación actual en Europa y España

La situación actual en Europa parece estar dominada por los clados caracterizados por la mutación D614G en la espícula del virus (clados G), de tal modo que alcanzan una frecuencia superior al 70% en países como Austria, Bélgica, Dinamarca, Finlandia, Francia, Alemania, Grecia, Islandia, Italia, Luxemburgo, Portugal, Escocia, Suecia y Suiza. Países como Inglaterra (64%), Holanda (63%), Gales (61%) y España (52.9%) presentan las tasas más bajas del continente.

Sin embargo, la frecuencia de la mutación L84S en la proteína NS8 en la mayoría de países europeos es menor del 1% (Bélgica, Dinamarca, Inglaterra, Finlandia, Francia, Italia, Noruega,

Suecia, Suiza y Gales). En otros países como Luxemburgo (2%), Portugal (3%), Islandia (3%), Alemania (3%), Escocia (4%) y Grecia (8%) esta frecuencia es algo mayor. España es el país europeo con mayor frecuencia de virus de esta clado, alcanzando un valor del 40% y sumando 3 de cada 5 virus de este grupo secuenciados en toda Europa; parece ser que estas variantes llegaron antes y en mayor intensidad a España y han conseguido mantenerse en el tiempo (es lo que se conoce como un efecto fundador).

En el caso de España, los 15 genomas secuenciados más antiguos proceden de muestras de Madrid, Valencia, Segovia y Granada de finales de febrero del 2020 y de Guadalajara, Burgos, Álava, Vizcaya, La Rioja, Tenerife y Orense de la primera semana de marzo [16]. De estas 15 secuencias, 9 pertenecen al clado 19B, 3 al clado 20A y una a los clusters 19A, 20B y 20C [16]. El hecho de que el clado 19B se haya encontrado en al menos 9 provincias diferentes hace pensar que se haya podido producir un efecto fundador que explicaría la alta prevalencia de dicho clado en España en contraposición a lo observado en otros países de nuestro entorno (Figura 3).

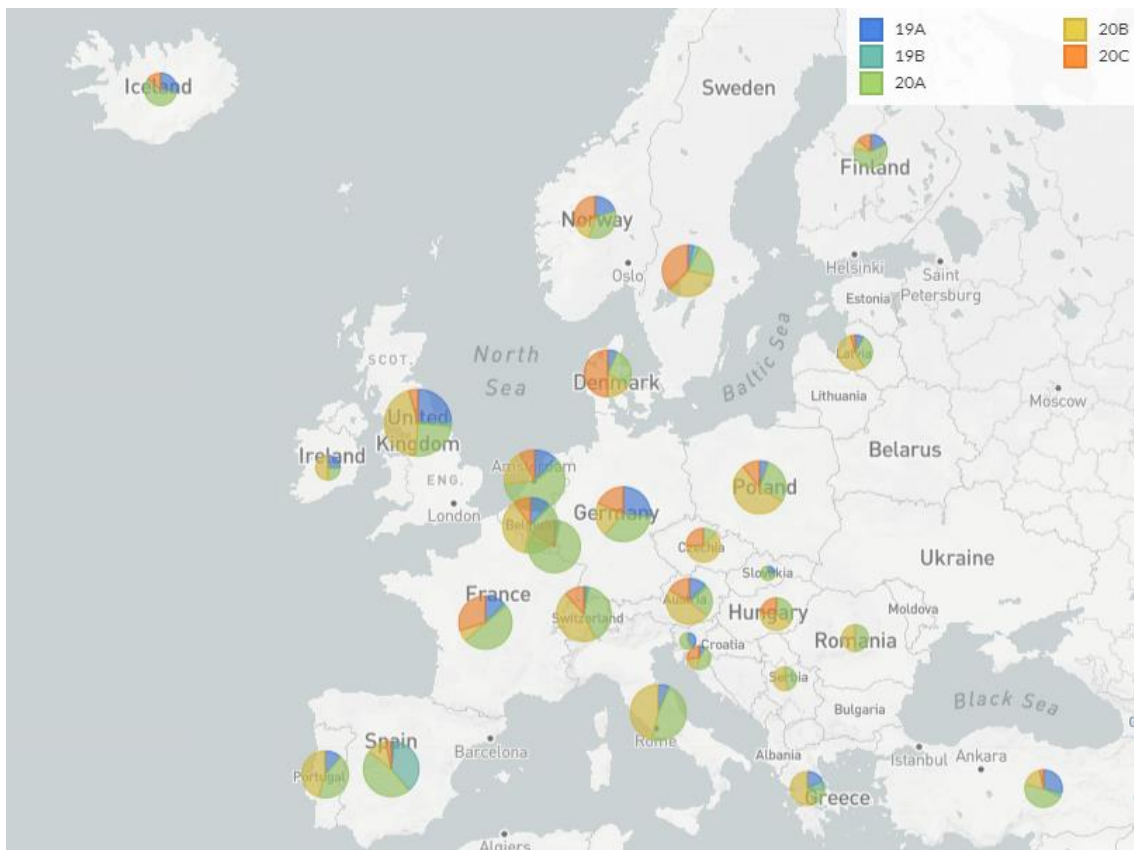


Figura 3. Frecuencia de los diferentes clados definidos por Nexstrain en Europa

Ejemplo: ¿cómo evoluciona el SARS-CoV-2 en un crucero?

Durante la pandemia de COVID-19, los cruceros se han convertido en modelos controlados para el estudio de la evolución del SARS-CoV-2, ya que debido a las cuarentenas impuestas tras

la detección de personas infectadas suponen una importante fuente de información sobre cómo se comporta el virus cuando entra en una población cerrada en la que no se producen flujos de entrada/salida de personas. El estudio de un brote en uno de estos cruceros en el que se secuenciaron 28 genomas completos ha revelado cómo el virus encontrado en 20 de esos 28 fue capaz de mutar hasta en 24 posiciones durante el periodo de cuarentena que duró 3 semanas. Las otras 8 secuencias fueron idénticas entre ellas y a la secuencia de referencia encontrada en el mercado de Wuhan de la que se supone que derivaron todas las demás [18].

Sin embargo, el caso del crucero *Diamond Princess*, que partió de Yokohama (Tokyo, Japón) el 20 de enero del 2020, es posiblemente el que mayor notoriedad ha adquirido debido a la gravedad del brote de SARS-CoV-2. En este crucero viajaban un total de 2.666 pasajeros y 1.045 miembros de la tripulación, que fueron obligados a permanecer en una cuarentena de dos semanas hasta el 19 de febrero de 2020 tras detectarse que un pasajero de 80 años había dado positivo a COVID-19 [19].

A todas las personas a bordo ($n=3.711$) se les realizó un test diagnóstico mediante RT-qPCR y hasta el día 8 de marzo del 2020; un total de 697 casos fueron detectados (18.8%) y 7 personas acabaron falleciendo. Tras secuenciar el genoma completo de 73 muestras de estos individuos infectados por SARS-CoV-2 se determinó la presencia de la mutación G11083T en todas ellas, por lo que los autores del estudio sugieren un único ancestro común para las 73 secuencias y por tanto una única introducción como causante del brote.

Un total de 29 de esas secuencias (40%) resultaron ser idénticas, lo que podría indicar un evento “supercontagiador” originado por el primer paciente detectado y antes de establecerse la cuarentena en el crucero. Otro 40% de esos genomas resultaron ser secuencias únicas que acumulaban entre 1 y 5 mutaciones con respecto a las secuencias de dicho cluster, lo que podría ser el resultado de pequeños eventos de diseminación ocurridos durante la cuarentena (ya que los autores pudieron trazar varios casos de contagios entre personas que compartían camarote) [19]. Por tanto, el caso del *Diamond Princess* supuso un importante experimento en el que se pudo comprobar cómo se puede transmitir el SARS-CoV-2 en una comunidad controlada.

¿Existe atenuación o incremento en la virulencia del SARS-CoV-2?

Uno de los genes más importantes del genoma del virus es el gen S, que codifica la espícula del virus. Ésta proteína media la entrada del virus en la célula diana mediante la unión de la subunidad S1 al receptor ACE2 celular y la posterior fusión de las membranas viral y celular a través de la subunidad S2. La mutación D614G alcanzó una prevalencia global superior al 40% en abril del 2020, haciendo que saltaran las alarmas sobre el posible significado biológico de dicha mutación en la que un aminoácido polar con una cadena lateral cargada (ácido aspártico) es sustituido por un pequeño aminoácido sin cadena lateral como la glicina.

[Según un estudio reciente](#), los virus que presentan la mutación D614G en el gen de la espícula empezaron a diseminarse en Europa a finales de enero-principios de febrero del 2020 y en el momento que se introduce en un nuevo territorio es capaz de convertirse en la forma dominante [20]. Los investigadores defienden que este hecho ha tenido lugar a nivel global y también en países como Inglaterra, Francia, Alemania, Italia, Holanda, Japón, EEUU o Australia, e incluso a nivel de ciudades como Nueva York o Washington.

Sin embargo, analizando los datos de los primeros 300 genomas completos de España, esta tendencia no parece cumplirse, ya que los virus G614 no han conseguido desplazar a los virus D614 y ambas formas genéticas parecen coexistir con frecuencias similares. Una posible explicación de este hecho en España podría ser un efecto fundador; es decir, es posible que los virus D614 (y especialmente los virus del clado 19B) llegarán antes y en mayor número a diferentes puntos de la geografía española. De hecho, entre los 15 genomas de España más antiguos (detectados durante la última semana de febrero y la primera de marzo de 2020), 10 de ellos eran D614 (67%) y 5 G614 (33%) [18].

Estos autores lanzan diferentes hipótesis sobre las posibles ventajas de los virus G614 en respecto a las variantes D614 que circulaban originalmente en la ciudad de Wuhan [20]. Las posibles ventajas estarían basadas en características estructurales y/o inmunológicas. Desde el punto de vista estructural, la mutación G614 podría actuar disminuyendo la interacción entre las subunidades S1 y S2, facilitando la liberación de S1 [20].

En este sentido, otro estudio ha analizado cómo se comportan ambas variantes integradas en un retrovirus pseudotipado en cultivos celulares de células HEK293T [21]. Estos autores han observado una mayor infectividad en los pseudovirus con la variante G614 y una mayor incorporación de la proteína S en el pseudovirión en comparación con la variante D614. Aunque la unión al receptor y su neutralización por plasma de individuos convalecientes fue equivalente, los autores concluyen que la mutación G614 parece que aporta estabilidad a la proteína S del virus y podría ser la causa de una transmisión más eficiente [21].

Desde el punto de vista inmunológico, el equivalente de la mutación G614 en SARS-CoV parece integrarse en el epítipo inmunodominante *LYQDVNC*, que es reconocido por anticuerpos aislados de pacientes recuperados tras la infección por SARS-CoV. [22]. Por tanto, esta mutación podría conferir resistencia frente a respuestas ‘protectoras’ mediadas por anticuerpos dirigidos frente a D614, provocando mayor susceptibilidad frente a la reinfección por virus G614 [20].

Un tercer mecanismo que podría estar implicado es un fenómeno denominado amplificación de la infección dependiente de anticuerpos (ADE), debido a que el epítipo donde se encuentra insertado es un epítipo ADE inmunodominante [22]. Aunque este tipo de fenómenos ADE han sido previamente observados en la infección por SARS-CoV, de momento no existen evidencias de que juegue ningún papel en el caso de SARS-CoV-2 [23].

En cuanto a la posible mayor virulencia de las variantes G614 tampoco parece haber demasiada evidencia científica hasta el momento. Existe un estudio reciente que ha relacionado la letalidad de la infección por SARS-CoV-2 con la prevalencia de la mutación G614 en diferentes países del mundo, entre ellos España [24]. Estos autores encontraron que la letalidad estimada correlaciona con la proporción de virus G614 ($p < 0.02$), aunque en el caso de España los datos parecen ser un *outlier* en dicha correlación [24]. Sin embargo, los sesgos que se han podido cometer tanto en el cálculo de la tasa de letalidad como de la prevalencia de la mutación G614 hacen que las conclusiones de este estudio sean poco robustas.

Por otra parte, en una cohorte de 453 individuos infectados por SARS-CoV-2 de Sheffield, Inglaterra, no se han observado diferencias en la severidad de la enfermedad causada por variantes G614 y D614, ya que las proporciones de pacientes infectados por ambas variantes



MINISTERIO
DE CIENCIA
E INNOVACIÓN



Instituto de Salud Carlos III

fueron similares tanto en pacientes ambulatorios, como en los que requirieron hospitalización o cuidados intensivos [20].

Conclusiones

En definitiva, la circulación de la variante G614 parece estar imponiéndose en muchos países del mundo en general y de Europa en particular. Las explicaciones de este hecho se centran principalmente en dos posibles hipótesis: ventaja selectiva de las variantes G614 frente a las D614 o un efecto fundador, donde la variante que llega antes y con mayor fuerza es la que acaba imponiéndose.

Unos pocos trabajos científicos, la mayoría publicados de momento sin revisión por pares, apuntan a una ventaja selectiva de los virus con la mutación G614; incluso se ha demostrado una mayor infectividad en ensayos de laboratorio. Sin embargo, estos resultados deben tomarse con cautela, ya que es necesario demostrar que estos resultados suponen realmente una ventaja en la transmisión del virus.

Por otro lado, el análisis de las más de 1600 importaciones del virus en Reino Unido ha demostrado que la variante mayoritaria G614 llegó antes y en mayor intensidad que las variantes D614. En el caso de España, el análisis de las variantes más antiguas detectadas hasta el momento (última semana de febrero y primera de marzo), indicaría un alto porcentaje de secuencias del clado 19B formado por variantes D614, que pudieron generar brotes simultáneos en al menos 9 provincias españolas. Este hecho podría ir a favor de un posible efecto fundador, y explicaría además la alta prevalencia del clado 19B observada en España en comparación con los demás países europeos.

En cualquier modo, ya sea por una ventaja evolutiva o por un efecto fundador o incluso por una combinación de ambas, la monitorización de los clusters de transmisión mediante epidemiología genómica es fundamental para conocer la diversidad genética del virus presente en un territorio concreto; evaluar su dispersión; estudiar si las medidas de protección utilizadas han sido exitosas; ayudar en el estudio de posibles nuevos brotes, y facilitar la toma de decisiones a la hora de imponer medidas de contención, entre otras cuestiones.

Madrid, 2 de julio de 2020

Informe realizado por Francisco Díez-Fuertes. Resumen divulgativo: José A. Plaza.

Grupo de Análisis Científico de Coronavirus del Instituto de Salud Carlos III.

Integran este grupo los Drs Mayte Coiras, Francisco Díez, Elena Primo, Cristina Bojo, Beatriz Pérez-Gómez, Francisco David Rodríguez, Esther García-Carpintero, Luis María Sánchez, José A. Plaza y Débora Álvarez. Está coordinado por el Dr José Alcamí.

BIBLIOGRAFÍA

1. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML, Zhang YL, Dai FH, Liu Y, Wang QM, Zheng JJ, Xu L, Holmes EC, Zhang YZ. A new coronavirus associated with human respiratory disease in China. *Nature* 2020; 579(7798):265-269.
2. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology* 2020, 5:536–544.
3. Shu, Y., McCauley, J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *EuroSurveillance* 2017; 22(13)
4. Duchene S, Featherstone L, Haritopoulou-Sinanidou M, Rambaut A, Lemey P, Baele G. Temporal signal and the phylodynamic threshold of SARS-CoV-2. *bioRxiv* 2020. doi: <https://doi.org/10.1101/2020.05.04.077735>
5. Zhao Z, Li H, Wu X, Zhong Y, Zhang K, Zhang YP, Boerwinkle E, Fu YX. Moderate mutation rate in the SARS coronavirus genome and its implications. *BMC Evol Biol.* 2004; 4:21
6. Cotten M, Watson SJ, Zumla AI, Makhdoom HQ, Palser AL, Ong SH, Al Rabeeah AA, Alhakeem RF, Assiri A, Al-Tawfiq JA, Albarrak A, Barry M, Shibl A, Alrabiah FA, Hajjar S, Balkhy HH, Flemban H, Rambaut A, Kellam P, Memish ZA. Spread, circulation, and evolution of the Middle East respiratory syndrome coronavirus. *mBio.* 2014; 5(1).
7. Li X, Zai J, Zhao Q, et al. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *J Med Virol.* 2020; 82(6): 602-611.
8. Andersen, K.G., Rambaut, A., Lipkin, W.I., Holmes, E.C., Garry, R.F. The proximal origin of SARS-CoV-2. (2020) *Nature Medicine* 2020, 26(4):450-452.
9. Benvenuto, D., Giovanetti, M., Salemi, M., Prosperi, M., De Flora, C. Junior Alcantara, L.C., Angeletti, S., Ciccozzi, M. The global spread of 2019-nCoV: a molecular evolutionary analysis. *Pathogens and Global Health* 2020, 114(2):64-67.
10. Xu, Y. Unveiling the Origin and Transmission of 2019-nCoV. *Trends in Microbiology* 2020, 28(4):239-240.
11. Böhmer MM, Buchholz U, Corman VM, Hoch M, Katz K, Marosevic DV, Böhm S, Woudenberg T, Ackermann N, Konrad R, Eberle U, Treis B, Dangel A, Bengs K, Fingerle V, Berger A, Hörmansdorfer S, Ippisch S, Wicklein B, Grahl A, Pörtner K, Müller N, Zeitlmann N, Boender TS, Cai W, Reich A, An der Heiden M, Rexroth U, Hamouda O, Schneider J, Veith T, Mühlemann B, Wölfel R, Antwerpen M, Walter M, Protzer U, Liebl B, Haas W, Sing A, Drosten C, Zapf A. Investigation of a COVID-19 outbreak in Germany resulting from a single travel-associated primary case: a case series. *Lancet Infect Dis.* 2020; S1473-3099(20):30314-5.
12. Hodcroft EB. Preliminary case report on the SARS-CoV-2 cluster in the UK, France, and Spain. *Swiss Med Wkly* 2020;150:9-10.
13. Oliver Pybus & Andrew Rambaut with Louis du Plessis, Alexander E Zarebski, Moritz U G Kraemer, Jayna Raghwan, Bernardo Gutiérrez, Verity Hill, John McCrone, Rachel Colquhoun, Ben Jackson, Áine O’Toole, Jordan Ashworth, on behalf of the COG-UK consortium. Preliminary analysis of SARS-CoV-2 importation & establishment of UK transmission lineages. 2020 <https://virological.org/t/preliminary-analysis-of-sars-cov-2-importation-establishment-of-uk-transmission-lineages/507>



14. Ana Da Silva Filipe, James Shepherd, Thomas Williams, Joseph Hughes, Elihu Aranday-Cortes, Patawee Asamaphan, Carlos Balcazar, Kirstyn Bruncker, Stephen Carmichael, Rebecca Dewar, Michael D Gallagher, Rory Gunson, Antonia Ho, Natasha Jesudason, Natasha Johnson, E. Carol McWilliam Leitch, Kathy Li, Alasdair MacLean, Daniel Mair, Sarah E. McDonald, Martin McHugh, Jenna Nichols, Marc Niebel, Kyriaki Nomikou, Richard J. Orton, Aine O'Toole, Massimo Palmarini, Yasmin A. Parr, Andrew Rambaut, Stefan Rooke, Sharif Shaaban, Rajiv Shah, Joshua B. Singer, Katherine Smollett, Igor Starinskij, Lily Tong, Vattipally B. Sreenu, Elizabeth Wastnedge, David L. Robertson, Matthew T.G. Holden, Kate Templeton, Emma C. Thomson. Genomic epidemiology of SARS-CoV-2 spread in Scotland highlights the role of European travel in COVID-19 emergence. medRxiv 2020. doi: <https://doi.org/10.1101/2020.06.08.20124834>.
15. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P, Bedford T, Neher RA. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 2018; 34(23):4121-4123.
16. Elbe S, and Buckland-Merrett G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Global Challenges* 2017;1:33-46.
17. Rambaut A, Holmes EC, Hill V, O'Toole A, McCrone JT, Ruis C, Plessis L, Pybus OG. A dynamic nomenclature proposal for SARS-CoV-2 to assist genomic epidemiology. bioRxiv 2020. doi: <https://doi.org/10.1101/2020.04.17.046086>.
18. Yeh TY & Contreras GP. Faster de novo mutation of SARS-CoV-2 in shipboard quarantine. [Preprint]. *Bull World Health Organ*. E-pub: 6 April 2020. doi: <http://dx.doi.org/10.2471/BLT.20.255752>
19. Sekizuka T, Itokawa K, Kageyama T, Saito S, Takayama I, Asanuma H, Naganori N, Tanaka R, Hashino M, Takahashi T, Kamiya H, Yamagishi T, Kakimoto K, Suzuki M, Hasegawa H, Wakita T, Kuroda M. Haplotype networks of SARS-CoV-2 infections in the Diamond Princess cruise ship outbreak. medRxiv 2020, doi: <https://doi.org/10.1101/2020.03.23.20041970>
20. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, Foley B, Giorgi EE, Bhattacharya T, Parker MD, Partridge DG, Evans CM, Freeman TM, de Silva TI, on behalf of the Sheffield COVID-19 Genomics Group, LaBranche CC, Montefiori DC. Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2. bioRxiv 2020. doi: <https://doi.org/10.1101/2020.04.29.069054>
21. Zhang L, Jackson CB, Mou H, Ojha A, Rangarajan ES, Izard T, Farzan M, Choe H. The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. bioRxiv 2020 doi: <https://doi.org/10.1101/2020.06.12.148726>
22. Wang, Q., Zhang, L., Kuwahara, K., Li, L., Liu, Z., Li, T., Zhu, H., Liu, J., Xu, Y., Xie, J., et al. (2016). Immunodominant SARS Coronavirus Epitopes in Humans Elicited both Enhancing and Neutralizing Effects on Infection in Non-human Primates. *ACS Infectious Diseases* 2, 361–376.
23. Jaume M, Yip MS, Cheung CY, Leung HL, Li PH, Kien F, et al. Anti-severe acute respiratory syndrome coronavirus spike antibodies trigger infection of human immune cells via a pH- and cysteine protease-independent FcγR pathway. *J Virol* 2011, 85:10582-10597.
24. Becerra-Flores M, Cardozo T. SARS-CoV-2 viral spike G614 mutation exhibits higher case fatality rate. *The International Journal of Clinical Practice* 2020. <https://doi.org/10.1111/ijcp.13525>

Otras referencias

- ¿Qué es el genoma? En la revista Méthode, Universidad de Valencia. <https://metode.es/revistas-metode/monograficos/que-es-el-genoma-2.html>
- ¿Qué es la secuenciación genética? En el Instituto de Salud Carlos III (ISCIII). <https://www.isciii.es/InformacionCiudadanos/DivulgacionCulturaCientifica/DivulgacionISCIII/Paginas/Divulgacion/InformeCoronavirusSecuenciacion.aspx>
- Informe sobre secuenciación genética. En la web CoNprueba, publicado por el Grupo de Análisis Científico sobre Coronavirus del Instituto de Salud Carlos III (ISCIII): <https://www.conprueba.es/secuenciacion-genetica-que-es-y-para-que-sirve>
- La genética traza el mapa de la dispersión mundial del virus. En Agencia SINC: <https://www.agenciasinc.es/Noticias/La-genetica-traza-el-mapa-de-la-dispersion-mundial-del-virus>
- Los genetistas consideran improbable que el coronavirus circulara masivamente por España antes de febrero. En Agencia SINC: <https://www.agenciasinc.es/Reportajes/Los-genetistas-ven-improbable-que-el-coronavirus-circulara-masivamente-en-Espana-antes-de-febrero>
- Nace la primera herramienta para rastrear el genoma del coronavirus en España. En Next Ciencia (Voz Populi). <https://www.efesalud.com/coronavirus-genetica-wuhan/>
- Qué información proporciona la secuenciación del genoma del coronavirus SARS-CoV-2. En Genotipia: https://genotipia.com/genetica_medica_news/sars-cov/