

FACULTAD DE CIENCIAS BIOLÓGICAS Y AMBIENTALES

**SARS-CoV-2: ANÁLISIS COMPARATIVO
DEL GENOMA VÍRICO DURANTE 2020 Y
2021 EN CASTILLA Y LEÓN, MADRID Y
CATALUÑA.**

**SARS-COV-2: COMPARATIVE ANALYSIS
OF VIRAL GENOME THROUGHOUT 2020
AND 2021 IN CASTILLA Y LEÓN,
MADRID AND CATALONIA.**

Autor: Alonso Capellán Sanjuán

Tutores: Luis E. Sáenz de Miera y Camino Fierro

GRADO EN BIOLOGIA

Julio, 2022

ÍNDICE

1. INTRODUCCIÓN	1
1.1. CARACTERÍSTICAS DEL SARS-CoV-2.	1
1.1.1. Clasificación y morfología del SARS-CoV-2.	1
1.1.2. Genes y proteínas del SARS-CoV-2.	2
1.2. Objetivos	6
2. MATERIALES Y MÉTODOS.....	7
2.1. Recogida de datos y procedimiento.....	7
2.2. Herramientas informáticas.	7
2.3. Tipo de datos.	8
3. RESULTADOS	9
3.1. Número de secuencias totales.	9
3.2. Porcentaje de mutantes, tasa de mutación, concentración y razón NS/SY	9
3.3. Selección en el genoma vírico, razón dN/dS.	13
3.4. Transiciones, transversiones, delecciones e inserciones relevantes	14
4. DISCUSIÓN.	16
5. CONCLUSIONES	24
6. MATERIAL SUPLEMENTARIO	25
7. REFERENCIAS.....	25
ANEXOS	31

ABREVIATURAS

%INS: Porcentaje de inserciones.

M.B.M.: Mutaciones por mutante.

%M.: Porcentaje de mutantes.

M.R.: Tasa de mutación.

%RDEL: Porcentaje de delecciones.

NS/SY: Proporción de mutaciones no sinónimas por cada mutación sinónima.

%RTS: Porcentaje de transiciones.

NSP: Proteína no estructural.

aa: Aminoácido.

NTD: Dominio N-terminal.

CC. AA.: Comunidades autónomas.

RdRp: Polimerasa dependiente de RNA.

cDNA: DNA complementario.

RTC: Complejo de replicación-transcripción.

CTs: Ciclos de amplificación mínimos.

RT-PCR: Reacción en cadena de la polimerasa con transcriptasa inversa.

DMVs: Vesículas de doble membrana.

dN/dS: Selección genética.

SNP: Polimorfismo de un solo nucleótido.

RESUMEN:

En este estudio se han analizado y comparado secuencias del genoma de *SARS-CoV-2* entre 2020 y 2021 provenientes de Castilla y León, Madrid y Cataluña dentro de la plataforma GISAID. Los resultados indican que la tasa de mutación media se incrementa hasta 5 veces en el periodo estudiado; Castilla y León tiene los valores más altos, mientras que Madrid tiene los más bajos. El gen *orf8* es el que más tasa de mutación tiene y más tendencia tiene a cambiar los aminoácidos para los que codifica, mientras que los genes *orf6*, *e* y *m* son los que menos cambian. El porcentaje medio de transiciones siempre supera a las transversiones y delecciones en las tres comunidades autónomas. La proporción entre mutaciones no sinónimas y sinónimas se ha mantenido relativamente constante entre todas las comunidades. La selección media del genoma (dN/dS) es purificadora en 2020 y positiva en 2021. Las mutaciones predominantes en ambos años son: C3037T (*orf1ab*: F106F), C14408T (*orf1ab*: P323L) y A23403G (*spike*: D614G).

Palabras clave: *SARS-CoV-2*, tasa de mutación, transversiones, transiciones, delecciones, dN/dS

ABSTRACT:

In this study, *SARS-CoV-2* genome sequences from Castilla y Leon, Madrid and Catalonia have been analyzed and compared between 2020 and 2021 within the GISAID platform. The results indicate that the average mutation rate increases up to 5 times within the period; Castilla y León has the highest values, while Madrid has the lowest. The *orf8* gene has the highest mutation rate and tendency to change its encoded amino acids, while the *orf6*, *e* and *m* genes are the ones that change the least. The average percentage of transitions in the three autonomous communities always exceeds transversions and deletions. Non-synonymous to synonymous ratio has been maintained constant in all sequences and time periods. The genome average selection (dN/dS) is purifying in 2020 and positive in 2021. The most common mutations on both periods are: C3037T (*orf1ab*: F106F), C14408T (*orf1ab*: P323L) and A23403G (*spike*: D614G).

Keywords: *SARS-CoV-2*, mutation rate, transversions, transitions, deletions, dN/dS

1. INTRODUCCIÓN

Desde el primer caso de COVID-19 registrado en la ciudad de Wuhan (provincia de Hubei, República Popular China) el 1 de diciembre de 2019, el virus que la causa, *SARS-CoV-2*, ha sufrido cambios tanto en su genoma como en la sintomatología asociada a la enfermedad. La sintomatología depende tanto de las características propias de la persona infectada, como de las del virus, siendo estas últimas de gran interés para su erradicación ya que gran parte de los tratamientos antivirales van dirigidos contra proteínas víricas, por ejemplo: “*Remdesivir*” inhibe la RNA polimerasa codificada en *orflab* (Gordon et al., 2020), aunque ya se han catalogado mutaciones relacionadas con la capacidad del virus para sobrevivir al fármaco como C16466T (P5401L) (Szemiel et al., 2021; Sarkar et al., 2021). Es de gran utilidad saber qué genes son los que tienen una mayor proporción de mutaciones, qué tipo de selección tienen, cuál es el carácter de las mutaciones encontradas, cuáles son las más relevantes y su implicación funcional respecto al proceso infeccioso. Todo esto ayudaría en el hallazgo de estrategias de lucha contra futuras pandemias.

1.1. CARACTERÍSTICAS DEL SARS-CoV-2

1.1.1. Clasificación y morfología del SARS-CoV-2.

El *SARS-CoV-2* es un virus RNA (reino Orthornavirae), monocatenario positivo (filo Pisuviricota), por lo tanto, se clasifica dentro del Grupo IV de la clasificación de Baltimore. Este grupo traduce directamente su material genético en los ribosomas de la célula huésped cuando la infectan, a diferencia de los virus RNA monocatenarios negativos. El *SARS-CoV-2* codifica una proteasa **M^{PRO}** característica propia de la clase Pisoviricetes. De entre los órdenes encontrados en esta clase, este virus está clasificado dentro de los Nidovirales debido a que sus genes estructurales están separados de los genes no estructurales (por marcos de lectura abiertos) y además tiene un deslizamiento ribosómico +1 en *orflab* que altera la traducción en los ribosomas (Kelly et al., 2020).

El virus de morfología globular y rodeado de una membrana de bicapa fosfolipídica, posee proteínas **SPIKE** con forma de maza ancladas a la membrana (**Figura 1**) dando un aspecto de “corona solar” a microscopía electrónica de transmisión, esta morfología es característica de la familia Coronaviridae. Dentro de esta familia los géneros más importantes son: *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* y *Deltacoronavirus*; los dos

primeros provienen del viroma de los murciélagos y los dos últimos del viroma de los cerdos y aves. El SARS-CoV-2 es un *Betacoronavirus* humano, que, a diferencia de otros como el *HCoV-OC4*, carece de la proteína hemaglutinina esterasa (HE) (Vijgen et al., 2005).

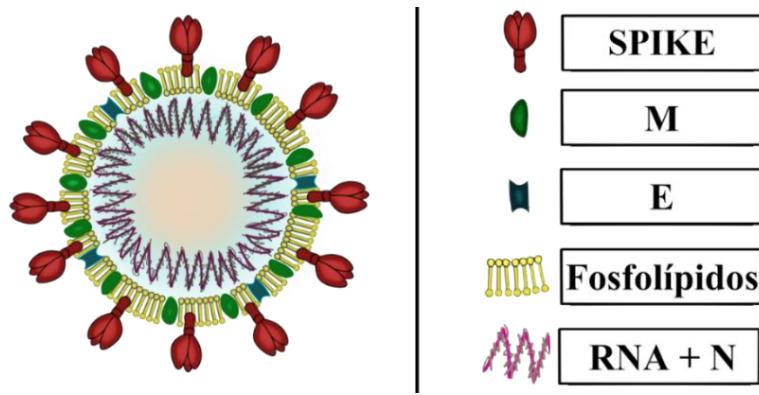


Figura 1. Estructura (virión) de SARS-CoV-2. Se resaltan las proteínas estructurales y la membrana formada por una bicapa fosfolipídica. Esquema elaborado a partir de Lim, I. et al. (2021).

1.1.2. Genes y proteínas del SARS-CoV-2.

El genoma del SARS-CoV-2 se caracteriza por tener 29891 nucleótidos totales en longitud (**Figura 2**) y 12 genes reconocidos, que ordenados por tamaño en orden decreciente serían: *orf1ab* (21289nt), *s (spike)* (3821nt), *n* (3821nt), *orf3a* (827nt), *m* (668nt), *orf3b* (455nt), *orf7a* y *orf8* (365nt), *e* (227nt), *orf6* (185nt), *orf7b* (131nt) y *orf10* (116nt). El 85% del genoma está ocupado por *orf1ab* y *s*. Por último, tenemos las regiones 5'-UTR (265nt) y 3'-UTR (216nt), ambas secuencias están implicadas en la formación de la estructura tallo-bucle que confiere estabilidad y protección al genoma de RNA viral, y es esencial para la replicación de este.

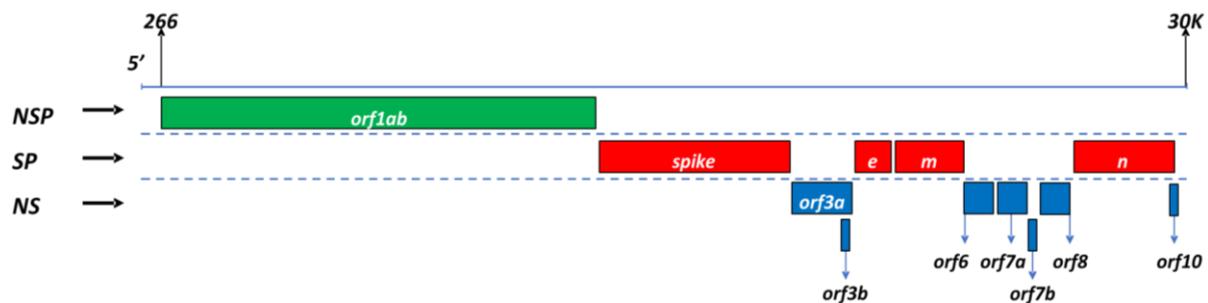


Figura 2. Genoma de SARS-CoV-2. Se puede observar la distribución de los genes en función de la posición. NSP (proteínas no estructurales), SP (proteínas estructurales) y NS (proteínas accesorias). Esquema elaborado a partir de Gordon, D.E et al. (2020).

El virus más parecido genéticamente a la primera secuencia de SARS-CoV-2 (WIV04) es *BatCoV-RaTG13*, siendo este último un betacoronavirus aislado del murciélagos *Rhinolophus affinis*, endémico de la región de Yunnan, China; por esto se propone que SARS-CoV-2 pudo

haberse generado por un salto zoonótico desde este, hasta el ser humano, aunque otras hipótesis especulan sobre un posible origen artificial (Sallard et al., 2021; Coccia, M., 2022). No está del todo claro el origen de SARS-CoV-2 pero existen evidencias de su capacidad para recombinarse de forma natural (Haddad et al., 2021) y, además, durante el 1º cuatrimestre de 2022 aparecieron las variantes XE, XF y XD que provienen de recombinaciones: la primera proviene de las variantes ómicron BA.1 y BA.2, y las 2 últimas provienen de variantes derivadas del linaje B.1.617 (Delta/Kappa) con variantes de ómicron. (Mohapatra et al., 2022).

Existen 3 tipos de proteínas codificadas en el virus: proteínas estructurales, proteínas no estructurales (NSP) y proteínas accesorias (NS). Las proteínas estructurales son aquellas que envuelven al virus y están codificadas por los genes: *s*, *e*, *m* y *n*. Las proteínas accesorias son aquellas que involucran funciones secundarias: juegan papeles implicados en la evasión inmunitaria o interaccionando con otras proteínas, estas son codificadas por los genes: *orf3a*, *orf3b*, *orf6*, *orf7a*, *orf7b*, *orf8* y *orf10*. Las proteínas no estructurales (NSP) están involucradas en la replicación y transcripción del genoma viral y se encuentran formando parte de las poliproteínas **PP1a** y **PP1ab**, ambas codificadas por el gen *orf1ab* y en conjunto se componen de 7097 aminoácidos. El final de **PP1a** y el comienzo de **PP1ab** está localizado en la posición nucleotídica 13176-13177 del *orf1ab* (aminoácido número 4394) y no finaliza en un codón de parada, por lo tanto, ambas poliproteínas se solapan y no están separadas. Ambas poliproteínas (**Figura 3**) están compuestas por 16 proteínas no estructurales (NSP), de entre las cuales las más importantes son la **NSP3 (PL^{PRO})**, **NSP5 (M^{PRO})** y la **NSP12 (RdRp)**, siendo esta última la encargada de replicar el genoma vírico (Gao et al., 2020). **M^{PRO}** es una proteasa que se autoescinde de la poliproteína **PP1a** al formarse en el ribosoma y posteriormente provoca cortes en los puntos que delimitan las proteínas **NSP4** y desde **NSP6** hasta **NSP16**, dejándolas libres y funcionales (Malone et al., 2022). Cuando **NSP4** se ha liberado, **PL^{PRO}** escinde **NSP1** y **NSP2**, además provoca la formación de DMVs (junto con **NSP4** y **NSP6**) cuya función es la de albergar el RTC. Las proteínas **NSP7** a **NSP16** se asocian entre sí formando el RTC que se introduce en las distintas DMVs y se encarga de producir la transcripción y replicación del material genético del virus (Malone et al., 2022).

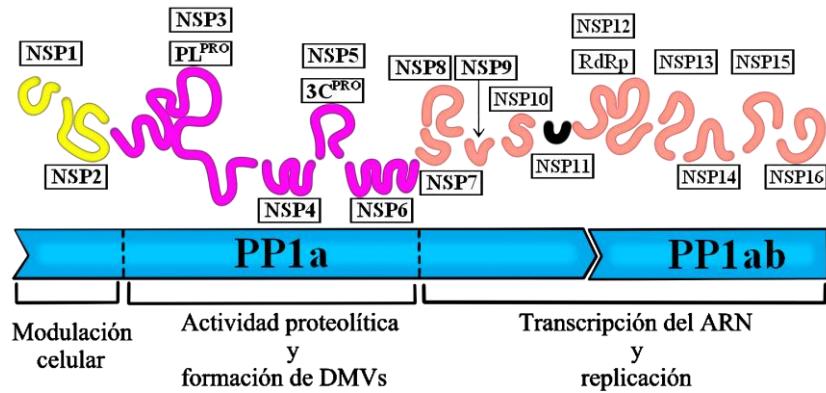


Figura 3. Poliproteínas codificadas en el gen *orf1ab* de SARS-CoV-2. Las poliproteínas **PP1a** y **PP1ab** se acompañan de los nombres de las proteínas que las integran (NSP) y en la parte inferior se nombran las funciones asociadas a dichas proteínas. Esquema elaborado a partir de ViralZone (2020).

La segunda proteína más grande es **SPIKE** (gen *s*; 1273 aminoácidos) se caracteriza por ser una única proteína estructural con 2 dominios: S1 y S2. El dominio S1 consta de 2 regiones: NTD (dominio N-terminal) y RBD (dominio de unión al receptor), esta última se encarga de unirse al receptor **ACE2** de las células humanas, mientras que S2 tiene los péptidos que intervienen en la fusión del virus con la membrana celular y proceder a la infección (Hoffmann et al., 2020). La proteína **SPIKE** necesita 2 enzimas localizadas en la membrana celular con distinta actividad proteasa (**Figura 4**) para que la unión a **ACE2** funcione correctamente: la furina celular y la **TMPRSS2**; la primera se une a la región que existe entre S1 y S2 (Lan et al., 2020; Premkumar et al., 2020) y la segunda a la región S2' (Peacock et al., 2021); una vez que esto ocurre, S1 se separa de S2 y el virus puede entrar en la célula por endocitosis (vesícula), o por fusión (solo se difunde el material genético del virus). Existe una vía de entrada alternativa mediada por los receptores de membrana celular **NRP1** y **NRP2** (neuropilinas) que tienen la capacidad de unirse al motivo Arg-Arg-Ala-Arg (RRAR) entre la posición (aa) 682 y 685 de la proteína **SPIKE** (**Figura 5**), actuando de forma similar a la furina. (Daly et al., 2020)

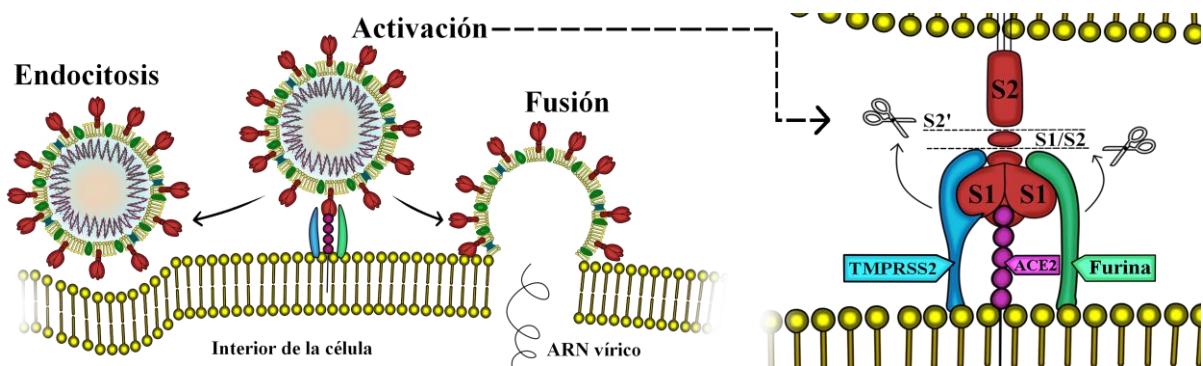


Figura 4. Proceso de unión entre SPIKE y ACE2. A la izquierda se muestran las 2 vías de entrada a la célula que puede seguir el virus cuando la unión entre **S1** y **ACE2** se ha realizado. A la derecha se detalla el proceso de activación. Esquema elaborado a partir de Lippi et al. (2020).

Una de las diferencias fundamentales entre *SARS-CoV-2* y el *SARS-CoV* original es que el primero tiene la secuencia de aminoácidos P-R-R-A entre S1 y S2 de la proteína **SPIKE** (**Figura 5**) que es la región del corte por la furina celular. (Wrobel et al., 2020)

Species/Abbrv	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	
1. MN996528.1 SARS-CoV-2 isolate WIV04 complete	I	G	A	G	I	C	A	S	Y	Q	T	Q	T	N	S	P	R	R	A	S	
2. MN996532.2 Bat coronavirus RaTG13 complete	I	G	A	G	I	C	A	S	Y	Q	T	Q	T	N	S	-	-	-	R	S	V
3. AY278488.2 SARS coronavirus BJ01 complete	I	G	A	G	I	C	A	S	Y	H	T	V	S	L	-	-	-	R	S	T	

Figura 5. Alineamiento parcial de la proteína SPIKE de *SARS-CoV-2* de Wuhan (*WIV04*), el coronavirus del murciélagos de herradura (*RaTG13*) y el *SARS-CoV* original (*BJ01*). La secuencia de aminoácidos marcada en rojo (P-R-R-A) forman parte del corte por furina. En la columna izquierda se muestran los nombres de las especies de coronavirus.

La proteína **NS3a** (gen *orf3a*; 275 aminoácidos) posee funciones distintas y poco conocidas; en cultivos celulares *in vitro* produce apoptosis (Ren et al., 2020), pero es probable que juegue un papel importante bloqueando la formación de autolisomas celulares (Ghosh et al., 2020; Miao et al., 2021) por lo tanto, es una proteína esencial para la supervivencia del virus cuando se encuentra dentro de la célula. La proteína **NS3b** (gen *orf3b*; 151 aminoácidos) se caracteriza por funcionar como un antagonista de interferones; es decir, inhibe mecanismos de inmunidad innata o celular en células infectadas (Konno et al., 2020). El gen *orf3b* está codificado dentro de *orf3a* (aunque dentro de este se consideran también los genes *orf3c* y *orf3d*), por lo tanto, la región que abarca la totalidad de *orf3a* codifica también 3 sub-genes distintos (Redondo et al., 2021). La proteína estructural **E** (gen *e*; 75 aminoácidos) (*envelope small protein*) es similar a la del *SARS-CoV* de 2003, su función principal consiste en formar poros en las células infectadas (Nieto-Torres et al., 2014), además juega un papel importante en el ensamblaje del virus en el aparato de Golgi (Nieto-Torres et al., 2015); también se sabe que interactúa con las proteínas que participan en las uniones celulares estrechas como **PALS1** debido a que **E** presenta el motivo peptídico D-L-L-V en la región terminal (Teoh et al., 2010). La proteína estructural **M** (gen *m*; 222 aminoácidos) (*multi-pass membrane protein*) está implicada en facilitar el ensamblaje del virus en el aparato de Golgi de la célula hospedadora mediante la interacción con otras proteínas víricas (Boson et al., 2021). La proteína **NS6** (gen *orf6*; 61 aminoácidos) interrumpe la formación del complejo celular de importación nuclear al unir carioferina- α -2 y carioferina- β -1, provocando el bloqueo de la expresión de genes que codifican productos antivirales (Miorin et al., 2020); también inhibe la producción celular de **IFN- β** mediante el bloqueo de la translocación nuclear del interferón **IRF3** (Xia et al. 2020). La proteína **NS7a** (gen *orf7a*; 121 aminoácidos) juega un papel importante como antagonista de la teherina (**BST2**), provocando la supresión de su actividad antiviral (Taylor et al., 2015). La proteína **NS7b** (gen *orf7b*; 43 aminoácidos) cuya

función es en gran parte desconocida, induce la apoptosis en cultivos celulares (Yang et al., 2021). La proteína **NS8** (gen *orf8*; 121 aminoácidos) juega un papel importante en la regulación inmunitaria: se une al receptor **IL17A** provocando respuestas proinflamatorias (Chan et al., 2020) por lo tanto, puede estar relacionada con la severidad de la enfermedad; se sabe que el gen *orf8* proviene de una duplicación de *orf7a*, por lo tanto, ambos son genes parálogos (Neches et al., 2021). La proteína estructural **N** (gen *n*; 419 aminoácidos) (*nucleoprotein*) se encarga de almacenar el material genético del virus en estado de virión y también está implicada en la replicación viral (Stertz et al., 2007); el gen *n* codifica 2 subgenes: *orf9b* (proteína **NS9b**) y *orf9c* (proteína **NS9c**), el primero antagoniza la interacción del translocador mitocondrial **TOM70** con **HSP90** debilitando la cascada de señalización para la expresión de IFN de tipo 1 (Gao et al., 2021), el segundo interactúa con los receptores sigma y con la vía NF-kB (Redondo et al., 2021). La proteína **NS10** (gen *orf10*; 38 aminoácidos) es un péptido cuya función es desconocida, pero se sabe que posee actividad inmunogénica (Mishra, 2020), aunque, la amplificación del gen *orf10* se utiliza extensivamente para detectar el virus mediante PCR en diagnóstico clínico (Hassan et al., 2020), otros autores defienden la eliminación de este gen ya que hay algunas evidencias de su carácter no codificante (Pancer et al., 2020).

1.2. Objetivos

Los objetivos principales de este trabajo consisten en:

- A- Comparar los valores de las tasas de mutación, concentración de mutaciones, cantidad de mutaciones no sinónimas por cada sinónima (NS/SY), transiciones, transversiones, delecciones e inserciones entre las comunidades de Castilla y León, Madrid y Cataluña durante 2020 y 2021.
- B- Comparar los valores de las tasas de mutación, concentración de mutaciones, cantidad de mutaciones no sinónimas por cada sinónima (NS/SY), razón Ti/Tv, delecciones y tipo de selección (dN/dS) en cada gen del virus entre 2020 y 2021.
- C- Encontrar, catalogar y explicar los efectos de las mutaciones relevantes más importantes en cada uno de los genes desde 2020 a 2021 presentes en las comunidades estudiadas.

2. MATERIALES Y MÉTODOS

2.1. Recogida de datos y procedimiento.

Se recogieron el mayor número de secuencias posibles del genoma de *SARS-CoV-2* desde la plataforma web GISAID (<https://www.gisaid.org/>) con una proporción de ‘N’ (errores de lectura en la secuencia) menor o igual al 5% y procedentes de las CC.AA. de Castilla y León, Madrid, Andalucía y Cataluña de 2020 y 2021. En las secuencias analizadas se trabaja con cDNA que es equivalente al RNA vírico. El año 2021 se distribuyó en 3 cuatrimestres (2021-Q1, Q2 y Q3). Las secuencias se separaron y analizaron por cada uno de los genes del virus. Andalucía se omitió en la estadística por la escasez de secuencias en 2021-Q3, pero se registraron sus respectivas mutaciones y se anotaron las más comunes. Se adoptaron estos criterios para evitar la mayor cantidad de errores posibles.

2.2. Herramientas informáticas.

El software utilizado para realizar los alineamientos de las secuencias con menos de 100 muestras han sido Mega 7.0.26 (Kumar et al., 2015) y Mega 11 (Kumar et al., 2020); para los alineamientos con más de 100 muestras se ha utilizado el servicio web ofrecido por el EBI (Instituto Europeo de Bioinformática), en ambos casos se ha utilizado la herramienta Clustal Omega (Sievers et al., 2018). Para exportar las secuencias a un formato tabulado separado por comas (.csv) se ha utilizado Ugene (Okonechnikov et al., 2021); la razón entre sustituciones sinónimas y no sinónimas se ha llevado a cabo con DNAsp (Rozas et al., 2017), las tablas, figuras y los porcentajes incluidos en los resultados se elaboraron en Excel (ver. 2105). Además, se han escrito y ejecutado scripts en Perl (*The Perl Foundation*, 2021) para convertir caracteres con ambigüedad (Nassar et al., 2017) de la IUPAC para nucleótidos (K, R, Y, W, S, M, H, B, V, D) a caracteres ‘N’ dentro de las secuencias. También se ha utilizado la herramienta online MEME (*Mixed Model Effects Model of Evolution*) (Murrell et al., 2012) desde la página web Datamonkey para calcular la relación dN/dS exacta de cada gen (valor ω). Para comparar los valores de las tasas de mutación, concentración de las mutaciones, razón NS/SY y razón Ti/Tv entre los genes estructurales y no estructurales (incluido *orf1ab*) en ambos años, se ha utilizado el software PAST 3.24 (Hammer et al., 2001) mediante el método estadístico ANOVA para comparar aquellos con valores de distribución normal y la prueba de Kruskal-Wallis para aquellos con valores de distribución no normal.

2.3. Tipo de datos.

Se trataron por separado el conjunto de todas las mutaciones y el conjunto de aquellas que aparecen al menos un 5% sobre la misma posición de las secuencias analizadas, a estas las llamamos mutaciones relevantes. Los datos expresados en todas las mutaciones son: el porcentaje de mutantes totales (%M.); la tasa de mutación (*Mutation Rate*, M.R.) que se define por la cantidad de mutaciones totales sobre la suma de la cantidad de nucleótidos de cada una de las secuencias para un intervalo de tiempo, que nos indica cuántas mutaciones adquieren las secuencias seleccionadas por unidad de tiempo; la selección genética (dN/dS, también llamada *Ka/Ks*) se calcula a través del cociente de las mutaciones no sinónimas entre los sitios no sinónimos dividido por el cociente de las mutaciones sinónimas entre los sitios sinónimos y el valor resultante nos indica qué genes han adquirido mutaciones que facilitan (selección positiva, $\omega > 1$), dificultan (selección purificadora, $\omega < 1$) o no provocan ninguna alteración (selección neutra, $\omega \approx 1$) en la supervivencia y reproducción del virus (Miyata et al., 1980; Nei et al., 1986); la concentración de las mutaciones es un valor que nos indica el grado de dispersión de las mutaciones en una secuencia (valores bajos, mayor dispersión; valores altos, mayor concentración) y se calcula por el cociente de la tasa de mutación (en porcentaje) relativo al porcentaje de posiciones con mutaciones; las mutaciones no sinónimas (NS) son aquellas que producen un cambio en el aminoácido original respectivo a la posición y las mutaciones sinónimas (SY) son todas aquellas mutaciones que no provocan un cambio, la razón o proporción NS/SY indica cuántas mutaciones no sinónimas hay por cada mutación sinónima, es semejante a dN/dS pero no se considera la diferencia en el número de sitios sinónimos y no sinónimos, y finalmente, las mutaciones por mutante (*Mutations By Mutant*, M.B.M.) que son el total de nucleótidos de una o varias secuencias multiplicado por la tasa de mutación (M.R.). Los datos analizados sobre las mutaciones relevantes son: porcentaje de mutaciones de tipo transición (%RTS), cuyo valor expresa el porcentaje de mutaciones en las que el nucleótido (pirimidina o purina) se sustituye por otro del mismo tipo; por el contrario, el porcentaje de mutaciones de tipo transversión (%RTV) cuantifica las mutaciones en la que la naturaleza del nucleótido sustituido es distinta; el porcentaje de mutaciones de tipo delección (%RDEL) e inserción (%INS) se calculan a partir de la cantidad de delecciones e inserciones de forma respectiva; la suma de los valores de %RTS, %RTV, %RDEL e %INS es siempre 100. Finalmente, la razón Ti/Tv que no es más que la cantidad de transiciones entre transversiones, este valor nos da la cantidad de transiciones que hay por cada transversión; no es una variable igual de importante que la dN/dS o que la M.R., pero, algunos autores sugieren

que un valor superior de transversiones frente al de transiciones significa una reducción de la eficacia biológica en virus RNA (Lyons et al., 2017) que indica si cada partícula de virus empeora la capacidad de multiplicarse en una célula.

3. RESULTADOS

3.1. Número de secuencias totales.

Las secuencias recogidas en la base de datos del GISAID se muestran en la **Tabla 1**. Las CC. AA. de Castilla y León, Madrid y Cataluña suman 4269 secuencias. Las secuencias de Andalucía se omitieron de la estadística, pero se incluyeron en la lista de mutaciones relevantes.

Tabla 1. Número de secuencias encontradas de cada comunidad en 2020 y 2021. En la última fila se calcula el número total de secuencias por cada intervalo temporal.

Origen	Secuencias 2020	Secuencias 2021-Q1	Secuencias 2021-Q2	Secuencias 2021-Q3
Andalucía	520	364	35	0
Castilla y León	289	119	506	364
Cataluña	123	257	201	1122
Madrid	480	122	338	348
Sec. suma	1505	942	1129	1834
Total		5188		

3.2. Porcentaje de mutantes, tasa de mutación, concentración y razón NS/SY.

El 100% de las secuencias muestreadas son mutantes comparadas con el virus original de Wuhan (**Tabla 2**), esto se debe a que *orflab* tenía al menos una mutación cuando llegó a España en 2020. Todos los genes incrementan su %M en ambos períodos a excepción de *orf10*. El gen *e* tiene los valores más bajos medios, *n* es el único gen que llega al 100% de mutantes en el 1^{er} y 2º cuatrimestre de 2021 y disminuye en el 3^{er} cuatrimestre de 2021.

Tabla 2. Porcentaje de mutantes encontrados (%M) de cada gen de SARS-CoV-2 en 2020 y 2021. Se muestran 2 filas por cada intervalo de muestreo, la primera representa las medias y la segunda las desviaciones entre las CC. AA. (σ). Valores en porcentaje.

	<i>orflab</i>	<i>spike</i>	<i>orf3a</i>	<i>orf3b</i>	<i>e</i>	<i>m</i>	<i>orf6</i>	<i>orf7a</i>	<i>orf7b</i>	<i>orf8</i>	<i>n</i>	<i>orf10</i>
2020	100	91	23	15	2	28	4	2	2	23	41	25
σ	0	8	13	0.3	1	15	1	1	1	14	26	24
2021-Q1	100	100	40	25	3	59	4	20	2	78	100	46
σ	0	0	7	9	1.3	18	2	7	0.6	7	0	19
2021-Q2	100	100	69	22	3	64	7	59	44	97	100	2
σ	0	0	21	8	1	25	3	24	28	7	0	1
2021-Q3	100	100	100	28	3	99	5	96	94	98	77	4
σ	0	0	0	9	1	1	1	4	3	2	20	1
\bar{x} (2021)	100	100	70	25	3	74	5	58	47	91	92	17
σ	0	0	9	9	1	15	2	12	11	5	7	7

Respecto a la tasa de mutación (M.R.) (**Tabla 3**), en 2020 los genes, ordenados con los valores de forma creciente son: *orf7a* (6.7E-5), *e* (8E-5), *orf7b* (1.7E-4), *orf1ab* (2.5E-4), *orf6* (3.2E-4), *orf3a* (3.3E-4), *s* (3.7E-4), *orf3b* (4.1E-4), *m* (4.4E-4), *n* (5.2E-4), *orf8* (8.6E-4) y *orf10* (2.2E-3); mientras que en la totalidad del 2021, son: *e* (1.4E-4), *orf6* (2.91E-4), *orf10* (7E-4), *orf3b* (7.25E-4), *orf1ab* (9.1E-4), *m* (1.5E-3), *orf3a* (1.8E-3), *n* (3E-3), *s* (3.6E-3) *orf7a* (4.5E-3), *orf7b* (5.3E-3) y *orf8* (1.36E-2). Al observar la **Figura 6**, se puede comprobar que hay correlación total ($R^2 = 1$) de la tasa de mutación media del SARS-CoV-2 durante 2021. La tasa de mutación media entre 2020 y 2021 es de 9.5E-4 mutaciones/posición/año. No se ha encontrado diferencias significativas ($p > 0.05$) entre los genes estructurales y los no estructurales tanto en 2020 ($p = 0.55$) y 2021 ($p = 0.56$).

Tabla 3. Tasas de mutación (M.R.) de cada uno de los genes de SARS-CoV-2 en 2020 y 2021. Los valores de 2020 y 2021 están expresados en mutaciones/posición/año. Valores de los cuatrimestres de 2021 expresados en mutaciones/posición/cuatrimestre. Cada valor va acompañado de su desviación típica (σ).

	<i>orf1ab</i>	σ	<i>spike</i>	σ	<i>orf3a</i>	σ	<i>orf3b</i>	σ	<i>e</i>	σ	<i>m</i>	σ
2020	2.5E-4	2E-4	3.7E-4	2E-4	3.3E-4	2E-4	4.1E-4	2E-4	8.0E-5	2E-5	4.4E-4	3E-4
2021-Q1	8.4E-4	6E-4	2.8E-3	1E-3	7.3E-4	4E-4	6.6E-4	1E-4	1.6E-4	3E-5	8.4E-4	4E-4
2021-Q2	9.0E-4	4E-4	3.6E-3	1E-3	1.3E-3	7E-4	5.8E-4	1E-4	1.3E-4	7E-5	1.1E-3	4E-4
2021-Q3	9.3E-4	7E-4	3.8E-3	1E-3	2.3E-3	1E-3	8.3E-4	1E-4	1.4E-4	8E-5	1.9E-3	1E-3
\bar{x} (2021)	9.1E-4	6E-4	3.6E-3	1E-3	1.8E-3	8E-4	7.3E-4	1E-4	1.4E-4	6E-5	1.5E-3	7E-4
	<i>orf6</i>	σ	<i>orf7a</i>	σ	<i>orf7b</i>	σ	<i>orf8</i>	σ	<i>n</i>	σ	<i>orf10</i>	σ
2020	3.2E-4	2E-4	6.7E-5	2E-5	1.7E-4	4E-5	8.6E-4	1E-4	5.2E-4	3E-4	2.2E-3	1E-3
2021-Q1	1.9E-4	1E-4	5.8E-4	2E-4	1.8E-4	1E-4	5.9E-3	2E-3	3.9E-3	2E-3	3.0E-3	8E-4
2021-Q2	3.6E-4	1E-4	3.8E-3	1E-3	4.1E-3	9E-4	1.2E-2	6E-3	4.2E-3	2E-3	1.9E-4	1E-4
2021-Q3	2.8E-4	1E-4	6.0E-3	4E-3	7.4E-3	3E-3	1.7E-2	8E-3	2.0E-3	4E-4	3.6E-4	1E-4
\bar{x} (2021)	2.9E-4	1E-4	4.5E-3	2E-3	5.3E-3	1E-3	1.4E-2	5E-3	3.0E-3	1E-3	7.0E-4	3E-4

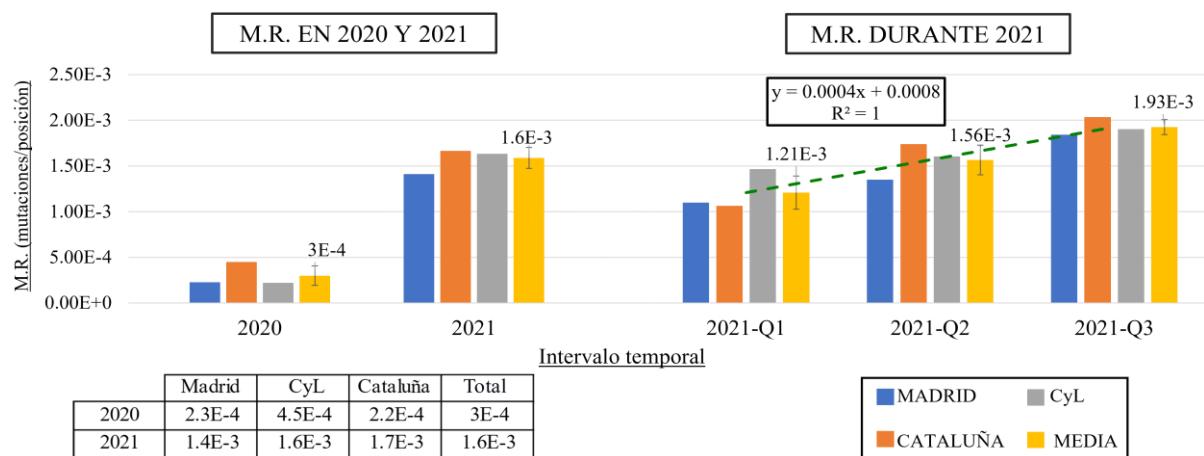


Figura 6. Tasas de mutación (M.R.) de SARS-CoV-2 en Castilla y León, Madrid y Cataluña en 2020 y 2021. La gráfica se compone del valor M.R. (entre todos los genes) en el eje vertical y los intervalos de tiempo en el eje horizontal. A la izquierda se comparan los valores (mutaciones/posición/año) en 2020 y el total de 2021, a la derecha se comparan los valores (mutaciones/posición/cuatrimestre) entre los 3 cuatrimestres de 2021; la ecuación de la recta se forma a partir de las medias de cada cuatrimestre. En la parte inferior izquierda se expresa en formato de tabla las tasas de mutación anuales medias en 2020 y 2021 por cada comunidad, en la columna final se expresan los valores promedio de cada año. El color amarillo indica la media de Castilla y León, Madrid y Cataluña.

La cantidad de mutaciones por mutante (M.B.M.) (**Figura 7**) se incrementa de 2020 a 2021. Al ser este valor otra forma de expresar la M.R., observamos que la correlación ($R^2 = 1$) es exactamente la misma entre los 3 cuatrimestres de 2021. La media entre 2020 y 2021 es de 28.04 mutaciones/mutante/año.

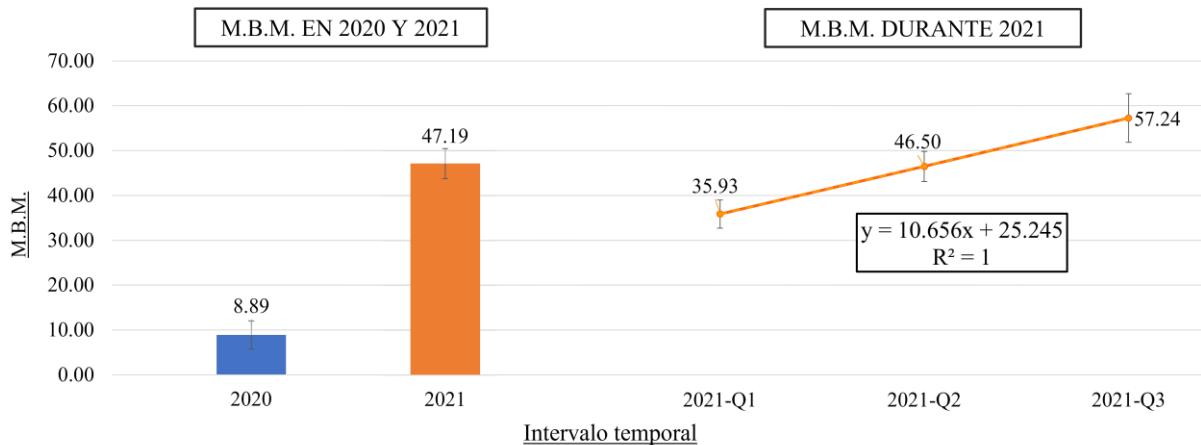


Figura 7. Mutaciones promedio por mutante (M.B.M) en el genoma de SARS-CoV-2 en 2020 y 2021. La gráfica tiene un eje vertical que indica el valor M.B.M. y el eje horizontal indica los intervalos de tiempo muestreados. A la izquierda se comparan los valores (mutaciones/mutante/año) de 2020 y el total de 2021, a la derecha se comparan los valores (mutaciones/mutante/cuatrimestre) de los 3 cuatrimestres de 2021; la ecuación de la recta se forma a partir de las medias de cada cuatrimestre.

Al ordenar por orden creciente los genes según la concentración de sus mutaciones, en 2020 tenemos: *orf7a* (3.8E-3), *orf7b* (3.9E-3), *e* (4.4E-3), *orf6* (4.6E-3), *orf3a* (0.011), *orf8* (0.014), *orf1ab* (0.02), *orf3b* (0.023), *n* (0.024), *s* (0.0256), *m* (0.028) y *orf10* (0.07); mientras que en la totalidad de 2021 son: *e* (6.14E-3), *orf6* (0.01), *orf3b* (0.013), *orf3a* (0.02), *orf7a* (0.035), *orf1ab* (0.038), *m* (0.04), *orf10* (0.044), *orf7b* (0.066), *n* (0.081), *s* (0.09) y *orf8* (0.134). No se ha encontrado diferencias significativas ($p > 0.05$) entre los genes estructurales y los no estructurales tanto en 2020 ($p = 0.889$) y 2021 ($p = 0.7$). Según se aprecia en la **Figura 8**, hay muy poca o ninguna correlación ($R^2 = 0.1892$) entre la concentración de las mutaciones en el genoma de SARS-CoV-2 a lo largo del 2021. El valor promedio de concentración entre 2020 y 2021 es de 0.03745.

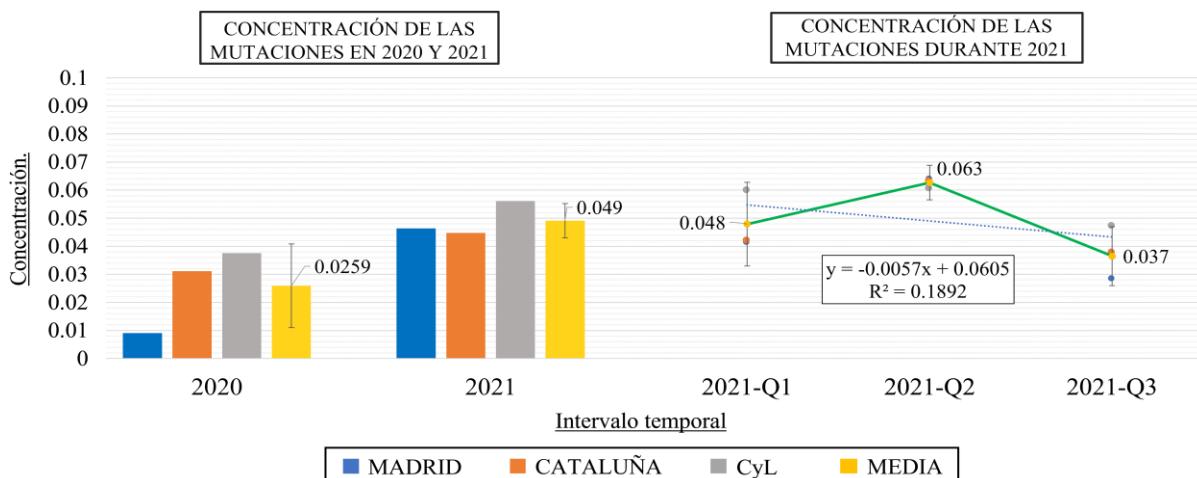


Figura 8. Concentración de las mutaciones promedio en el genoma de SARS-CoV-2 en 2020 y 2021. El gráfico muestra los valores de las concentraciones de las mutaciones entre todos los genes en el eje vertical, mientras que el eje horizontal está formado por los intervalos de tiempo muestreados. A la izquierda se comparan los valores de 2020 y el total de 2021, a la derecha se indican los valores de los 3 cuatrimestres de 2021; la ecuación de la recta se forma a partir de los valores de cada cuatrimestre.

Al calcular la razón NS/SY de las CC.AA. tanto en 2020 como en cada uno de los cuatrimestres de 2021 (**Figura 9**) encontramos siempre una tendencia constante de aproximarse a 1.5, en términos reales equivale a 3 mutaciones no sinónimas por cada mutación sinónima, es decir, hay aproximadamente un 60% de mutaciones no sinónimas y un 40% de mutaciones sinónimas tanto en 2020 como durante 2021. El promedio entre 2020 y 2021 es de 1.465. Al ordenar por orden creciente los genes según la relación NS/SY en 2020 tenemos: *m* (0.4), *orf6* (0.875), *orf1ab* (1.35), *orf7b = n = orf10* (1.5), *s* (1.65), *orf3a* (2.44), *orf8* (2.7), *orf3b* (2.833), *e* (3) y *orf7a* (4.33); mientras que, en la totalidad de 2021: *m* (0.56), *orf1ab* (1.17), *orf6* (1.46), *n* (1.56), *s* (1.9), *orf10* (2.625), *e* (2.66), *orf3a* (3.2), *orf3b* (3.33), *orf7a* (3.6), *orf8* (4.35) y *orf7b* (5.22). No se ha encontrado diferencias significativas ($p > 0.05$) entre los genes estructurales y los no estructurales tanto en 2020 ($p = 0.4$) y 2021 ($p = 0.085$).

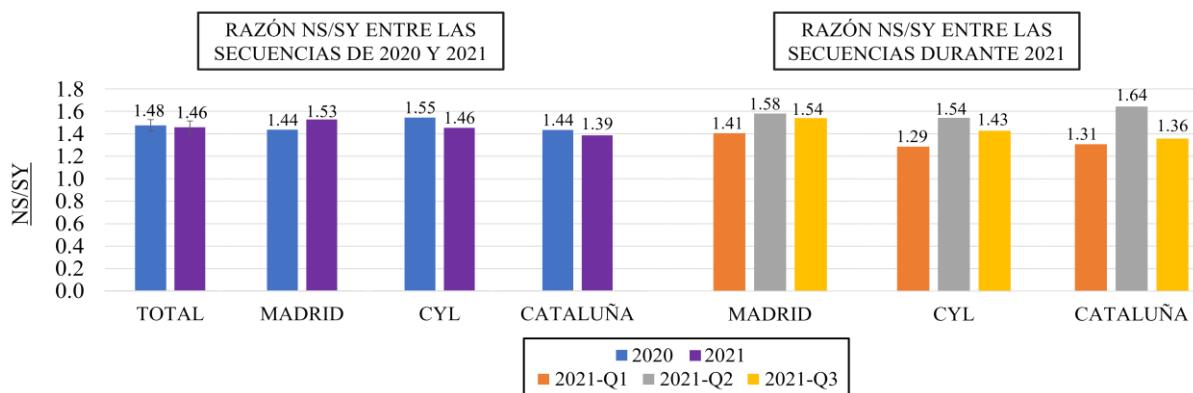


Figura 9. Razón de mutaciones no sinónimas entre sinónimas (NS/SY) de Madrid, Castilla y León y Cataluña en 2020 y 2021. El gráfico muestra el valor NS/SY en el eje vertical y en el eje horizontal encontramos las comunidades. A la izquierda se comparan los valores de 2020 y el 2021, a la derecha se comparan los valores de los 3 cuatrimestres de 2021.

3.3. Selección en el genoma vírico, razón dN/dS.

Respecto al valor de la relación dN/dS (**Tabla 4**), en 2020 los genes se ordenan en orden creciente de la siguiente forma: *m* (0.17), *orf6* (0.25), *orf7b* (0.41), *orf1ab* (0.6), *s* (0.68), *n* (0.88), *orf10* (1.02), *e* (1.024), *orf8* (1.03), *orf3a* (1.18) y *orf7a* (1.62) y *orf3b* (1.68); mientras que según los valores promedio de 2021, los genes se ordenan de la misma forma como: *m* (0.27), *orf1ab* (0.48), *orf6* (0.49), *e* (0.83), *s* (0.89), *n* (1.07), *orf10* (1.39), *orf3a* (1.43), *orf3b* (1.54), *orf8* (1.58), *orf7a* (2.38) y *orf7b* (2.65). No se ha encontrado diferencias significativas ($p > 0.05$) entre los genes estructurales y los no estructurales tanto en 2020 ($p = 0.34$) y 2021 ($p = 0.1$). Cuando observamos los datos gráficamente (**Figura 10**), no existe correlación ($R^2 = 0.0431$) durante 2021, pero, está claro que el genoma de SARS-CoV-2 pasa de tener una selección purificadora en 2020 (0.88) hasta una selección positiva en 2021 (1.25). El valor promedio entre 2020 y la totalidad de 2021 es de 1.065.

Tabla 4. Selección genética (dN/dS) de cada uno de los genes en 2020 y 2021. Los datos mostrados son la media de los valores ω (\bar{x}) seguido de la desviación típica (σ).

		<i>orf1ab</i>		<i>spike</i>		<i>orf3a</i>		<i>orf3b</i>		<i>e</i>		<i>m</i>	
		ω (\bar{x})	σ										
2020		0.60	0.05	0.68	0.20	1.18	0.30	1.69	0.57	1.02	0.33	0.17	0.06
2021-Q1		0.46	0.04	0.77	0.16	1.29	0.26	1.69	0.44	0.64	0.01	0.23	0.10
2021-Q2		0.47	0.02	1.13	0.39	1.77	0.84	1.55	0.60	0.68	0.24	0.30	0.15
2021-Q3		0.51	0.02	0.76	0.06	1.22	0.13	1.45	0.48	1.18	0.44	0.29	0.10
2021-\bar{x}		0.48	0.03	0.89	0.20	1.43	0.41	1.57	0.51	0.83	0.23	0.27	0.12
		<i>orf6</i>		<i>orf7a</i>		<i>orf7b</i>		<i>orf8</i>		<i>n</i>		<i>orf10</i>	
		ω (\bar{x})	σ										
2020		0.25	0.08	1.62	1.18	0.41	0.06	1.03	0.16	0.88	0.21	1.02	0.76
2021-Q1		0.41	0.15	3.44	2.40	2.09	1.30	0.97	0.11	1.14	0.14	0.97	0.54
2021-Q2		0.49	0.16	2.54	1.04	3.59	2.21	2.52	0.69	1.34	0.78	1.60	0.77
2021-Q3		0.59	0.30	1.15	0.24	2.27	1.23	1.26	0.46	0.73	0.20	1.61	1.20
2021-\bar{x}		0.49	0.21	2.38	1.23	2.65	1.58	1.58	0.43	1.07	0.38	1.39	0.84

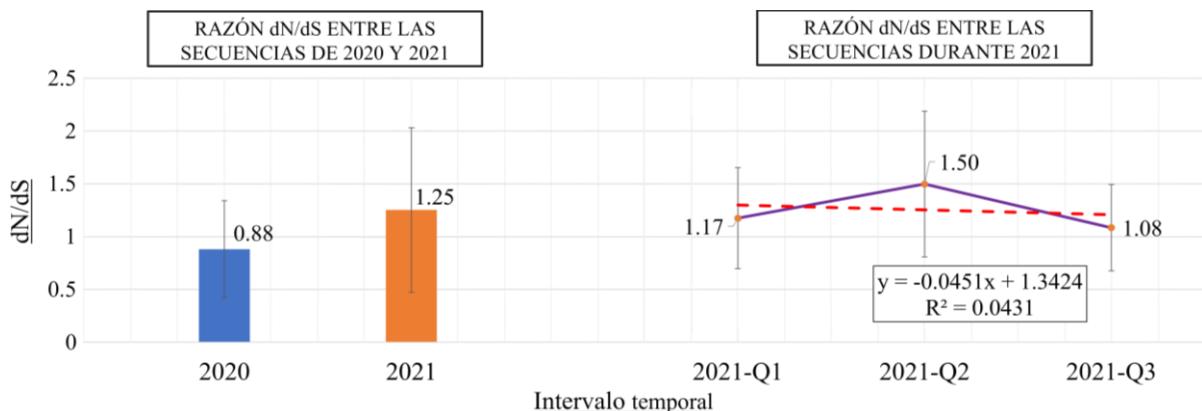


Figura 10. Selección promedio del genoma de SARS-CoV-2 en 2020 y 2021. El gráfico de puntos representa en el eje vertical el promedio de dN/dS entre todos los genes de SARS-CoV-2 y en el eje horizontal cada uno de los intervalos temporales muestreados. A la izquierda se comparan los valores de 2020 y el total de 2021, a la derecha se comparan los valores de los cuatrimestres de 2021; la ecuación de la recta se forma a partir de los valores de cada cuatrimestre.

3.4. Transiciones, transversiones, delecciones e inserciones en mutaciones relevantes.

La **Figura 11** y la **Figura 12** indican que, en 2020 los valores de %RTS y %RTV medios son de 81.5% y 17.9% respectivamente, mientras que en 2021 son de 50.6% y 22.8%. Entre 2020 y el total de 2021, los valores de %RTS y %RTV son de 66% y 20.5% de forma respectiva.

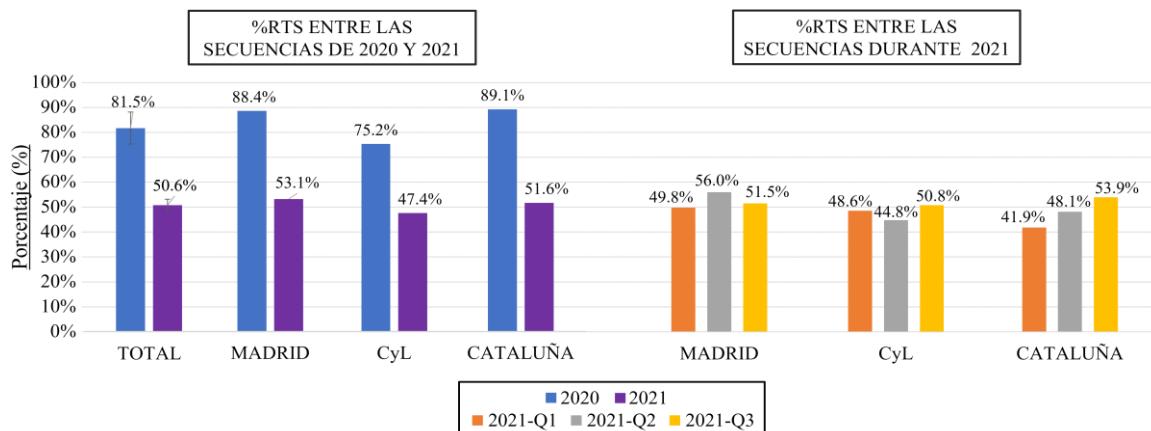


Figura 11. Porcentaje de transiciones en mutaciones relevantes de Madrid, Castilla y León y Cataluña en 2020 y 2021. El gráfico muestra en el eje vertical los porcentajes de transiciones en mutaciones relevantes (%RTS) y en el eje horizontal muestra las CC. AA. A la izquierda se comparan los valores de cada año, a la derecha se comparan los valores de los 3 cuatrimestres de 2021.

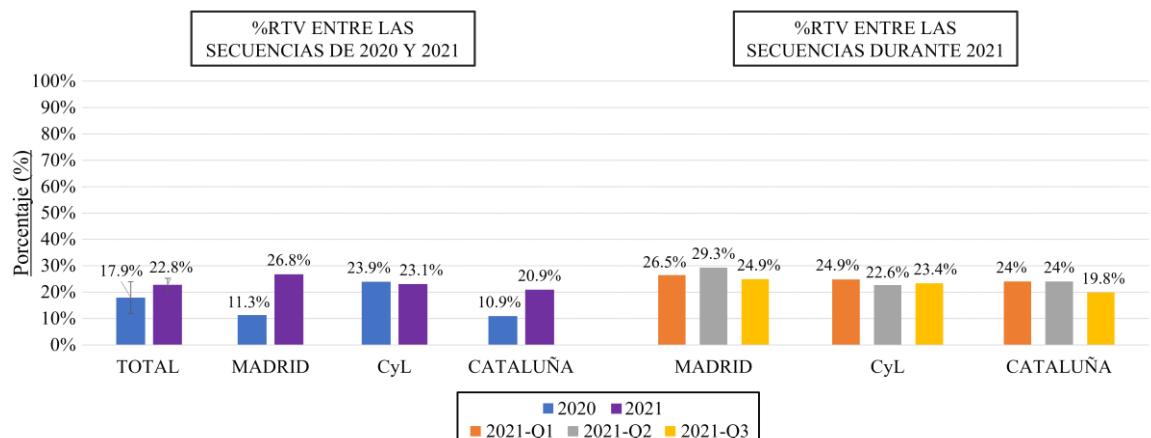


Figura 12. Porcentaje de transversiones en mutaciones relevantes de Madrid, Castilla y León y Cataluña en 2020 y 2021. El gráfico muestra en el eje vertical los porcentajes de transversiones en mutaciones relevantes (%RTV) y en el eje horizontal muestra las CC. AA. A la izquierda se comparan los valores de cada año, a la derecha se comparan los valores de los 3 cuatrimestres de 2021.

Los valores obtenidos en 2020 indican que solamente el gen *s* y *orf6* tienen delecciones entre las mutaciones relevantes, los porcentajes respectivos (%RDEL) de estos son 2.65% y 12.2%; mientras que, los genes ordenados de forma creciente según sus valores promedio durante 2021 son: *orf10* (0.23%), *n* (0.3%), *orf7a* (0.43%), *orf3a* (0.73%), *orf6* (1.33%), *orf3b* (7.29%), *orf1ab* (12.84%), *s* (43.14%) y *orf8* (78.54%). La **Figura 13** indica que el valor de las %RDEL en 2020 y 2021 son respectivamente 0.56% y 26.5%. El valor promedio entre 2020 y 2021 es de 13.4%. Respecto a las inserciones (%INS), al observar la **Figura 14**, identificamos que solamente aparecen en el 3º cuatrimestre de 2021; la media entre 2020 y 2021 es de 0.035%.

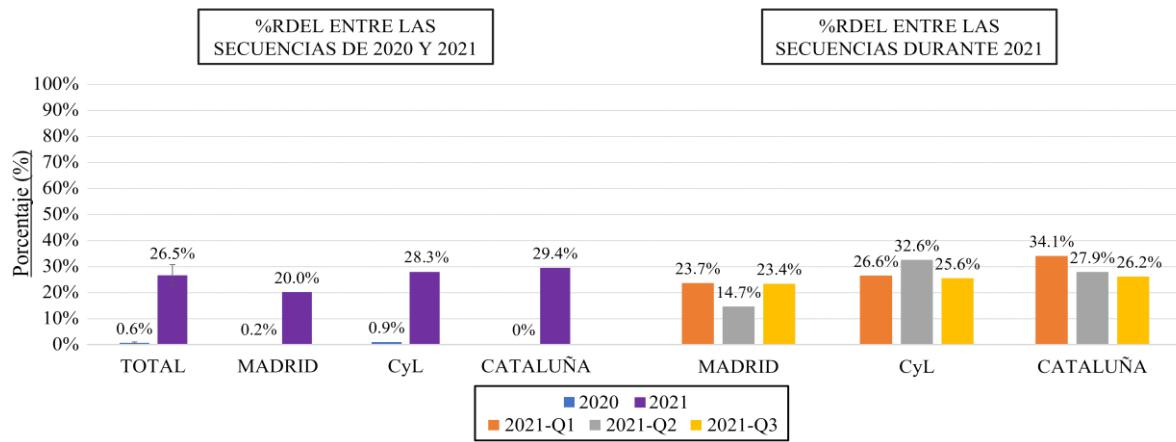


Figura 13. Porcentaje de delecciones en mutaciones relevantes (%RDEL) de Madrid, Castilla y León y Cataluña en 2020 y 2021. El gráfico de barras muestra en el eje vertical las delecciones en mutaciones relevantes (%RDEL) y las CC.AA. se muestran en el eje horizontal. A la izquierda se comparan los valores de 2020 y el total de 2021, a la derecha se comparan los valores de los 3 cuatrimestres de 2021.

Los datos obtenidos por orden creciente de los genes según la relación Ti/Tv en 2020 tenemos: *m* (0.16), *e* (0.2), *orf7b* (0.54), *orf3a* (0.76), *orf3b* (0.83), *orf10* (0.88), *orf7a* (1.75), *orf1ab* (3.44), *n* (3.44), *orf8* (7.87), *orf6* (11) y *s* (38.8); mientras que de media en 2021, los genes se ordenan: *orf10* (0.05), *m* (0.34), *orf3a* (0.78), *s* (0.8), *orf7b* (0.86), *orf7a* (0.87), *orf1ab* (0.95), *n* (1), *orf3b* (1.08), *orf8* (2.1), *orf6* (5.6) y *e* (6). No se ha encontrado diferencias significativas ($p > 0.05$) entre los genes estructurales y los no estructurales tanto en 2020 ($p = 0.67$) y 2021 ($p = 0.86$). Cuando calculamos la relación Ti/Tv de cada uno de los intervalos de tiempo, encontramos que: en 2020 el valor es de 4.56, en el 1º cuatrimestre de 2021 es de 1.82, en el 2º cuatrimestre de 2021 es de 1.96 y en el 3º cuatrimestre de 2021 es de 2.47; el total en 2021 es de 2.22. La media entre 2020 y 2021 es de 2.29.

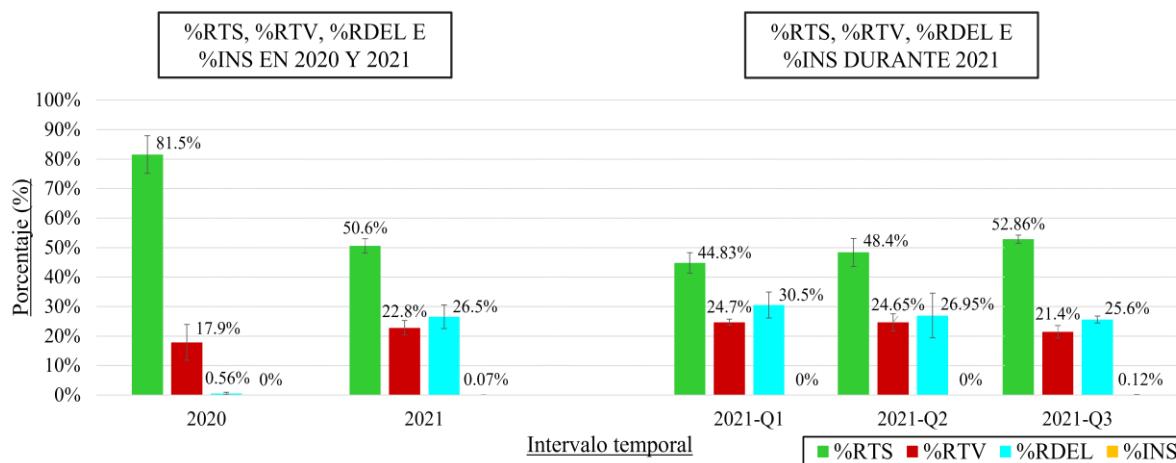


Figura 14. Porcentaje de transiciones, transversiones, delecciones e inserciones en mutaciones relevantes de SARS-CoV-2 en Madrid, Castilla y León y Cataluña en 2020 y 2021. El gráfico de barras representa los porcentajes respectivos a las variables %RTS, %RTV, %RDEL e %INS en el eje vertical, mientras que en el eje horizontal se representan los intervalos temporales. A la izquierda se comparan los valores de 2020 y el total de 2021, a la derecha se comparan los valores de los cuatrimestres de 2021.

4. DISCUSIÓN

El genoma de *SARS-CoV-2* empieza a cambiar de forma notable a partir del 1^{er} cuatrimestre de 2021 debido a que tanto la tasa de mutación, como la concentración de estas y el número de delecciones son mayores que en la totalidad del 2020; también, el virus pasa de tener una selección media purificadora a positiva justamente a partir del primer cuatrimestre de 2021. Una de las posibles causas por las que el virus comienza una tendencia a la selección positiva en el 2021 es debido a que las primeras vacunas que se empezaron a administrar desde diciembre de 2020 hayan favorecido en el virus aquellas mutaciones capaces de alterar los distintos antígenos principales localizados en S1, S2 y RBD que forman parte de **SPIKE** (Jagtapa et al., 2021), permitiéndole zafarse de los anticuerpos generados en aquellos vacunados; la mayor parte de las mutaciones relevantes destacadas encontradas en el gen *s* en el 3^{er} cuatrimestre de 2021 tienen relación con escapes inmunitarios (**ANEXO II**). Aún así, se puede observar que en el 2º y 3^{er} cuatrimestre de 2021 los valores ω en el gen *s* son menores de 1, por lo tanto, la tendencia a la selección positiva en el virus no solamente se explica por los cambios en **SPIKE** sino por la combinación de estos con mutaciones en otros genes que provocan sinergia entre sus efectos. Como ya hemos dicho, este estudio solo abarca a las comunidades de Castilla y León, Madrid y Cataluña; sería necesario extender el estudio al resto de España para tener resultados más concluyentes.

Se observa en todos los genes a excepción de *orf6* y *orf10* una tendencia al incremento de la tasa de mutación desde 2020 hasta finales de 2021 (**Tabla 3**), aunque curiosamente este último gen tiene el valor promedio más alto de todos en 2020, mientras que *orf8* lo tiene en 2021; por otra parte, *orf7a* y *e* tienen los valores promedio más bajos de 2020 y 2021 respectivamente. El aumento proporcional de cada gen entre 2020 y el promedio de 2021, en orden creciente son: *orf3b* (+68%), *e* (+77%), *m* (+193%), *orf1ab* (+252%), *orf3a* (+331%), *n* (+557%), *s* (+820%), *orf8* (+1241%), *orf7b* (+2224%) y *orf7a* (+5021%); por otro lado, los genes en los que ha disminuido la tasa de mutación de forma proporcional son: *orf6* (-13%) y *orf10* (-45%). Esto indica que hay diferencias con respecto a la velocidad con la que los genes adquieren o pierden mutaciones. Cuando comparamos los datos entre las 3 comunidades en 2020 (**Figura 6**) se puede observar que Castilla y León tiene la tasa de mutación más alta y Cataluña la más baja. Cataluña supera a las demás comunidades en el 1^{er} cuatrimestre de 2021, pero Castilla y León predomina en el 2º y 3^{er} cuatrimestre, por otra parte, Madrid tiene la tasa más baja. Si ordenamos por orden creciente las CC. AA. que entre 2020 y 2021 tienen mayores M.R.

tenemos: Madrid (8.15E-4), Cataluña (9.6E-4) y Castilla y León (1.03E-3). La tasa de mutación anual promedio entre las 3 comunidades se ha incrementado casi 5 veces desde 2020 (3E-4) a 2021 (1.6E-3).

Respecto a la concentración de las mutaciones (**Figura 8**), el valor se incrementa hasta 2.5 veces desde 2020 hasta 2021. Por otra parte, desde el 1º hasta el 2º cuatrimestre la concentración de las mutaciones aumenta, pero, los valores disminuyen en el 3º cuatrimestre; este fenómeno puede ser debido a que en el principio de diciembre de 2021 empieza a aparecer la variante ómicron BA.1 (Del Águila et al. 2022; Centro de Coordinación de Alertas y Emergencias Sanitarias, 2021), provocando que la abundancia de la variante delta en ese momento disminuya, es decir, las secuencias muestreadas que tenían en su momento mutaciones comunes de delta se agrupen junto con otras secuencias que tienen mutaciones en otras posiciones (ómicron) provocando un efecto de “dilución” de la variante dominante (delta). Los genes *orf10* y *orf8* son aquellos con los valores más altos en 2020 y 2021 respectivamente, mientras que, de igual forma, los genes *orf7a* y *e* tienen los valores más bajos. Cuando hacemos la media entre 2020 y 2021, los genes *orf8* (0.07462) y *e* (0.00525) tienen el valor más alto y bajo de forma respectiva. Todos los genes aumentan proporcionalmente la concentración de sus mutaciones desde 2020 a 2021 a excepción de *orf10* y *orf3b*, ordenados de forma creciente son: *orf3b* (-40%), *orf10* (-36%), *e* (+41%), *m* (+48.5%), *orf1ab* (+82%), *orf3a* (+88%), *orf6* (+120%), *n* (+240.7%), *s* (+254%), *orf7a* (+830%), *orf8* (+844%) y *orf7b* (+1631%).

Respecto a los tipos de mutaciones más comunes, predominan siempre las mutaciones no sinónimas respecto a las sinónimas. Al comparar los valores de la razón NS/SY (**Figura 8**) entre las distintas comunidades tanto en 2020 como en los 3 cuatrimestres de 2021, observamos que el valor tiende a 1.5, es decir, hay alrededor de 40% de mutaciones sinónimas y 60% de mutaciones no sinónimas en el genoma de *SARS-CoV-2* de forma constante. En lo que respecta a los genes: *orf7b* (3.36), *orf8* (3.52) y *orf7a* (3.97) aparecen con las razones NS/SY promedio más altas entre 2020 y 2021, mientras que, de igual modo, los genes: *orf1ab* (1.26), *orf6* (1.17) y *m* (0.4) siempre predominan entre los genes con mutaciones sinónimas. El gen *m* es el único con valores en ambos años por debajo de 1, por lo tanto, es un gen que tiende a conservarse.

Los genes de *SARS-CoV-2* que tienen selección purificadora (**Tabla 4**) en 2020 son: *orf1ab*, *s*, *m*, *orf6*, *orf7b* y *n*; mientras que en la totalidad de 2021 tenemos: *orf1ab*, *s*, *e*, *m* y *orf6*. Por otra parte, los genes con selección positiva en 2020 son: *orf3b*, *orf3a*, *orf8* y *orf7a*; mientras

que en la totalidad de 2021 tenemos: *orf3a*, *orf7a*, *orf7b*, *orf8*, *n*, *orf3b* y *orf10*. Los valores más bajos tanto en 2020 como en 2021 se encuentran en el gen *m*. Ningún gen evoluciona a valores por debajo de 1 en el 3º cuatrimestre de 2021 desde valores positivos en 2020, pero si hacemos la media de 2021 el gen *e* es el único que cambia de tener un comportamiento neutro-positivo en 2020 (1.023) a purificador en 2021 (0.834). Los valores más altos en 2020 y el 1º cuatrimestre de 2021 son encontrados en el gen *orf7a*, en los períodos restantes se encuentran en *orf7b*. Pasamos de 6 genes con selección positiva en 2020 a 7 en 2021. Los valores neutros ($\omega = 1$) suelen ser raros en los genes víricos, pero si nos fijamos en las medias totales de la **Figura 10** observamos que los valores del 3º cuatrimestre suelen acercarse al valor de 1. Los genes que tienden a la neutralidad los encontramos (**Tabla 4**) en 2020 y el 1º cuatrimestre de 2021, siendo estos: *e*, *orf10* y *orf8*. Al hacer la media entre ambos años, el gen *n* es el más cercano a la neutralidad (0.975). Los genes *s*, *orf3a*, *orf7b*, *orf8* y *n* presentan un crecimiento progresivo en su valor ω hasta el 2º cuatrimestre de 2021, pero sufren una posterior bajada pronunciada en el 3º cuatrimestre de 2021. Entre 2020 y 2021, aquellos genes que proporcionalmente han aumentado su dN/dS son por orden creciente son: *orf3a* (21%), *n* (22%), *s* (30%), *orf10* (37%), *orf7a* (46%), *orf8* (54%), *m* (56%) y *orf7b* (544%); por otra parte, aquellos que han disminuido son: *orf3b* (-8%), *e* (-18%) y *orf1ab* (-20%).

Al analizar el promedio de las mutaciones relevantes del genoma vírico (**Figura 14**); el valor de %RTS siempre es superior al %RTV y %RDEL. El valor de %RTS (**Figura 11**) en 2020 tiene valores que suelen aproximarse al 90% tanto en Madrid como en Cataluña, no así en Castilla y León cuyo valor es el más bajo (75.2%). A partir del primer cuatrimestre de 2021 los valores promedio de %RTS caen hasta un 30%, manteniéndose alrededor del 50% durante lo restante de 2021. Por otra parte, los valores de %RTV (**Figura 12**) de Madrid y Cataluña en 2020 suelen estar alrededor del 11%, excepto Castilla y León con un valor próximo al 24%.

El valor %RDEL (**Figura 13**) se incrementa significativamente en todas las CC. AA. a partir del 1º cuatrimestre de 2021, siendo el valor de 2020 completamente residual. Durante 2021, la Comunidad de Madrid tiene la menor proporción de delecciones, mientras que Cataluña tiene la mayor. Entre los genes: *orf6* tiene proporcionalmente más delecciones en 2020 que ningún otro, mientras que durante la totalidad de 2021 es *orf8* cuyo valor es muy alto (78.54%), por lo tanto, en este gen dominan las delecciones sobre todos los demás tipos de mutaciones, y es probable que la mutación 28248-28253del (*orf8*: 119-120 Δ) sea la responsable. Por último, las inserciones (%INS) empiezan a encontrarse desde noviembre de 2021 debido a la aparición de

algunas mutaciones como *orf3a*: ins104P. Al observar la relación Ti/Tv entre 2020 y 2021 de cada gen vemos que *s*, *orf8* y *orf6* siempre están entre los 3 de valores más altos, mientras que, por otra parte, el gen *m* suele ser aquel con valores más bajos.

Si miramos tanto los datos de la M.R., la relación dN/dS, la concentración de las mutaciones, la razón NS/SY y la razón Ti/Tv entre 2020 y 2021, el gen *orf8* tiene un comportamiento peculiar: siempre está entre los genes con los valores más altos, por lo tanto, es un gen que cambia mucho y concentra mucho sus mutaciones. También llaman la atención los genes *orf6*, *e* y *m*: el primero tiene valores muy bajos de M.R., concentración de las mutaciones, razón NS/SY y selección; el segundo siempre tiene los valores medios más bajos de la M.R. y concentración de las mutaciones entre la media de 2020 y 2021; el último tiene en la relación dN/dS y NS/SY más bajas en ambos años, además tiene los valores promedio de la razón Ti/Tv más bajos; todo esto indica que son genes que cambian muy poco debido a que: o adquieren muy pocas mutaciones, o estas son de carácter sinónimo. Los genes *e* y *m* además tienen también la característica adicional de no tener delecciones y de ser genes estructurales, este hecho puede tener el potencial para el desarrollo de nuevas vacunas o fármacos cuyo objetivo sea neutralizar o inhibir las proteínas que codifican. Además, se han encontrado anticuerpos anti-**M** (Jörrißen et al., 2021) y también anti-**E** (Szymczak et al., 2021) en suero convaleciente de personas recuperadas de la COVID-19.

No existen diferencias significativas de los parámetros de M.R., concentración, NS/SY, dN/dS y Ti/Tv entre los genes estructurales y no estructurales en ambos años, pero, en 2021, tanto los valores de NS/SY ($p = 0.085$) como los de dN/dS ($p = 0.1$) se acercan bastante al valor mínimo ($p = 0.05$) para rechazar la hipótesis nula (H^0 : no existen diferencias), por lo tanto, este estudio necesitaría un mayor número de replicas para poder confirmar con más exactitud estos resultados.

El gen *orf1ab* posee 4 mutaciones relevantes (**ANEXO II**) que se han mantenido desde 2020 hasta finales de 2021: C3037T (F924F), C4543T (T1426T), C14408T (P1415L) y C18877T (L6206L); las 2 primeras y la última son mutaciones sinónimas, la tercera se encuentra en la polimerasa, afecta a la replicación viral y se caracteriza por estar relacionada con la severidad de la enfermedad al combinarse con la mutación S: D614G (DemIr et al., 2021). Las mutaciones C3037T y C14408T aparecen un 98% de todas las secuencias muestreadas en 2020 y 20201, por lo tanto, son predominantes (**ANEXO III**). De entre las mutaciones más importantes encontradas en el 3^{er} cuatrimestre de 2021 tenemos: 686-694del (KSF141-143Δ),

G15451A (G571S) y C16466T (P5401L); la primera se sabe que afecta a la replicación y está relacionada con la reducción de la patogenicidad del virus (Benedetti et al., 2020) y la segunda se encuentra dentro del dominio “dedos” de la polimerasa vírica y produce un incremento a la tolerancia a fármacos antivirales inhibidores de la replicación como Remdesivir (Sarkar et al., 2021). Una de las mutaciones más importantes es 11288-11296del (SGF3675-3677Δ), esta aparece en algún momento del 1^{er} cuatrimestre de 2021 y provoca un aumento de la transmisibilidad del virus (Lee et al., 2021; Xia et al., 2020). Debe señalarse que, al codificar este gen una poliproteína y no una proteína única funcional, se tendría que analizar las secuencias que separan cada NSP (**Figura 3**) por separado como si fueran genes distintos para tener mucha mayor precisión en los resultados.

El gen *s* tiene 4 mutaciones (**ANEXO II**) que se han mantenido desde 2020 hasta finales de 2021: 21990-21993del (Y144Δ), C22227T (A222V), A23063T (N501Y) y A23403G (D614G); la primera forma parte de las variantes alfa (B.1.1.7) y ómicron BA.1 y BA.3 y su efecto es la aumentar la transmisibilidad (Sofia et al., 2022), la segunda es una mutación que se originó en España en verano de 2020 y produce una mayor unión del anclaje de RBD hacia los receptores ACE2 sin producir escapes inmunitarios (Ginex et al., 2021), la tercera está relacionada con un aumento en la capacidad de transmisión sin que haya mayor virulencia de la enfermedad y además provoca mayor afinidad hacia el receptor ACE2 (Public Health England, 2021), y por último; la mutación D614G es una mutación que se encuentra en todas las secuencias desde 2020 hasta finales de 2021, por lo tanto es predominante (**ANEXO III**), y está relacionada con menores CTs en RT-PCR, esto quiere decir que esta mutación confiere mayor aptitud biológica que el virus original (Baric, 2020) y produce una mayor carga viral, sobre todo en las regiones respiratorias superiores (Demír et al., 2020). Existen otras mutaciones que son importantes como 21765-21769del (HV69-70Δ) relacionada con infecciones en pacientes inmunodeprimidos (Public Health England, 2020) y es propia de la variante delta (B.1.617.2), alfa (B.1.1.7) y las variantes ómicron BA.4 y BA.5 (Patchsung et al., 2022; Ministerio de sanidad, 2022). Las mutaciones A23604A (P681H) y A23604G (P681R) son polimorfismos localizados en la región de unión a la furina celular (aa 681-688) y están relacionadas con una mayor capacidad de transmisión e infectividad (Public Health England, 2021), por último, la mutación G24914C (D1118H) la encontramos en S2, no está relacionada con incrementar la transmisibilidad o morbilidad, pero se cree que estabiliza la estructura y dinámica general de la proteína (Zhao et al., 2021). Entrando en el 3º cuatrimestre de 2021, las mutaciones C21618G (T19R), G21987A (G142D), 22029-22034del (EF156-

157Δ), T22917G (L452R) y C23525T (H655Y) están todas relacionadas con el escape inmunitario (Planas et al., 2021; Shen et al., 2021; Dussupt et al., 2021; Tchesnokova et al., 2021; Cao et al., 2022); la mutación L452R es común en las variantes delta (B.1.617.2), kappa (B.1.617.1), epsilon (B.1.427) y las variantes ómicron BA.4 y BA.5; la mutación H655Y es exclusiva de las variantes ómicron BA.1, BA.2, BA.3, BA.4 y BA.5 y tiene la capacidad de infectar gatos en condiciones de laboratorio (Braun et al., 2021). C22995A (T478K) es una mutación que aparece a finales del 2º cuatrimestre de 2021 y forma parte de la variante B.1.617.2 (delta) y está relacionada con el aumento de la virulencia de la enfermedad (Di Giacomo et al., 2021). Se puede decir que desde 2020 hasta el 2º cuatrimestre de 2021, las mutaciones relevantes del gen *s* están implicadas en su mayoría en mejorar el mecanismo de entrada a la célula hospedadora: por lo general la mayoría de ellas se concentran en S1 y son no sinónimas o delecciones, es decir, alteran la proteína, y por lo tanto están afectadas por la selección.

No se han encontrado mutaciones de *orf3a* que hayan perdurado desde 2020 hasta finales de 2021, pero la mutación C25626T (H78Y) (**ANEXO II**) está presente durante todo el 2021 y forma parte de la variante ómicron BA.2 (Desingu et al., 2022). Por otra parte, existen mutaciones como G25563T (Q57H) relacionadas con mayor infectividad del virus (Chu et al., 2021) y aparece desde las primeras secuencias de 2020. La mutación G25979T (G196V) tiene la capacidad de conferir estabilidad estructural al genoma viral (Bianchi et al., 2021). A partir del 3º cuatrimestre de 2021 encontramos la primera inserción más común: ins104P.

El gen *orf3b* tiene 2 mutaciones que han persistido desde 2020 hasta el final de 2021: G25906C (G48R) y G25996T (V78L). Las mutaciones representadas (**ANEXO II**) son de tipo no sinónimo a excepción de C25872T (Y36Y), y de entre todas las que tienen un impacto importante, C26060T (T99I) está relacionada con la desestabilización de la estructura del RNA viral (Hossain et al., 2021; Bianchi et al., 2021). Existen pocos estudios con relación al funcionamiento de este gen y mucho menos sobre sus mutaciones.

El gen de tipo estructural *e* (*envelope small protein*) es de pequeño tamaño en comparación con otros genes y además es el gen que como media, tiene menor tasa de mutación entre 2020 y 2021 (**Tabla 3**); por lo tanto, las mutaciones suelen ser bastante escasas. No hay ninguna mutación encontrada que haya persistido desde 2020 (**ANEXO II**), la única mutación relevante destacada registrada en 2020 es G26268T (E8D). Por otra parte, las mutaciones C26270T (T9I) y C26464T (L47L) aparecen en el 1^{er} cuatrimestre de 2021 y permanecen hasta finales del año;

la primera es la más importante y es propia de la variante ómicron BA.1 y se la relaciona con mayor estabilidad e integridad de la estructura del virus cuando este se ensambla en el aparato de Golgi (Mou et al., 2021). La mutación C26408T (S55F) aparece con gran frecuencia en el 2º cuatrimestre de 2021 en Madrid y Cataluña, el cambio que se produce es de aminoácido polar a no polar (serina por fenilalanina) y puede afectar al proceso de patogenicidad del virus (Mou et al., 2021). Las mutaciones G26428T (S62F) y C26444T (S68F) suelen ir acompañadas y se caracterizan respectivamente por mejorar el plegamiento de la proteína cuando se sintetiza en el ribosoma y mejorar la infectividad del virus (Mou et al., 2021).

El gen estructural *m* (*small membrane protein*) tiene los valores más bajos de NS/SY tanto en 2020 y 2021, por lo tanto, muchas de sus mutaciones son de carácter sinónimo. La cantidad de mutaciones relevantes catalogadas es baja (**ANEXO II**) en comparación con su tamaño. La única mutación que se encuentra tanto en 2020 como en el 3º cuatrimestre de 2021 es C26681T (F53F) de tipo transversión y sinónima. La mutación C26801G (L93L) es destacable porque aun siendo una mutación del tipo sinónima se la ha relacionado con una mayor severidad de la enfermedad (Xi et al., 2021). Por otro lado, existen 2 mutaciones no sinónimas que se mantienen presentes a lo largo de 2021 de forma continua: G2526T (A2S) y T26767C (I82T).

El gen *orf6* destaca por tener gran cantidad de mutaciones sinónimas relevantes (**ANEXO II**), pero también encontramos la aparición de mutaciones sin sentido (presencia de codones de terminación) como G27364T (E55*) y G27376T (E59*). La única mutación persistente desde 2020 hasta finales de 2021 es T27384C (D61L). Una mutación importante en el 3º cuatrimestre es A27259C (R20R) ya que solo aparece en la variante ómicron (BA.1) y algunas variantes descendientes de esta (Kailing et al., 2022)

El gen *orf7a* tiene 1 mutación relevante (**ANEXO II**) presente tanto en 2020 como a finales de 2021: G27506T (G38V). Por otro lado, mutación C27752T (T120I) se mantiene durante todo el 2021 y se caracteriza por ser una transición, con cambio de carga polar a no polar y está relacionada con reinfecciones e infecciones en vacunados (Chen et al., 2021), es una mutación relacionada con el escape inmunitario.

El gen *orf7b* tiene 1 mutación relevante (**ANEXO II**) presente tanto en 2020 como a finales de 2021: C27807T (L18L), de carácter sinónimo. Las mutaciones C27769T (S5L) y C27874T (T40I) son de tipo no sinónimo y en ambas se produce un cambio de carga de aminoácido polar a no polar tanto en S5L (serina por leucina) y en T40I (treonina por isoleucina).

El gen *orf8* destaca por tener 2 mutaciones relevantes (**ANEXO II**) encontradas tanto en 2020 y finales de 2021: C27944T (H17H) y G28086T (A65S); la primera es de carácter sinónimo y se ha comprobado que otorga estabilidad estructural al gen (Fang et al., 2021). Existen también, 2 mutaciones relevantes que son importantes en el 3^{er} cuatrimestre de 2021: C28093 (S67F) y 28248-28253del (119-120Δ); la primera provoca un cambio de aminoácido hidrofílico (serina) por otro de carácter hidrofóbico (fenilalanina) y la segunda es una doble delección codónica, ambas tienen la característica de alterar la afinidad de **ORF8** por el receptor celular **IL17RA** y por lo tanto modificar la respuesta inmunitaria (Hassan et al., 2021; Zinzula, 2020). Otras mutaciones encontradas son: A28111G (Y73C) y T28144C (L84S); la primera está relacionada con una transmisibilidad del 40% al 80% por encima del virus original (Public Health England, 2020) y la segunda está relacionada con mayor severidad y transmisibilidad de la enfermedad (Rehman et al., 2020). Las mutaciones C27972T (Q27*) y A28095T (K68*) son mutaciones sin sentido que provocarían péptidos recortados sin función biológica o con funciones no estudiadas.

El gen estructural *n* es el 3^{er} gen más grande en el genoma del *SARS-CoV-2*, por lo tanto, posee bastantes mutaciones. Existen 2 mutaciones relevantes (**ANEXO II**) que han perdurado durante todo el 2020 y 2021: GG28881-28882AA (R203K) y G28883C (G204R), ambas son contiguas en la secuencia del gen. Tanto GG28881-28882AA (R203K) como GG28881-28882TA (R203M) son polimorfismos con cambio nucleotídico doble, el primero es más prevalente en 2020 y el 1^{er} cuatrimestre de 2021, mientras que el segundo es más prevalente en el 2º y el 3º; ambos polimorfismos están relacionados con la mutación en *s* (D614G) formando parte de la llamada variante B1.617.2 (delta) cuya transmisibilidad es superior al virus originario de Wuhan (DemIr et al., 2020; Leary et al., 2020). Por otra parte, G28883C (G204R) está relacionada con la activación de cascadas de citoquinas y cuyo efecto es letal en la COVID-19 (Zhao et al., 2022). La mutación GAT28280-28282CTA (D3L) es de tipo codón completo y se caracteriza por un cambio aminoacídico de tipo ácido a no polar (ácido aspártico por leucina) y afecta al péptido señal de la proteína. La mutación A28461G (D63G) puede estar relacionada con un aumento en la transmisibilidad por aerosoles (Wassenar et al., 2022).

Por último, el extraño gen *orf10* de pequeño tamaño, funciones desconocidas, pero alta dN/dS (**ANEXO II**), tiene pocas mutaciones y pueden variar mucho incluso entre 2 intervalos temporales seguidos, pese a esto hay 3 mutaciones que persisten durante 2020 y 2021: G29402T (D377Y), C29614T (C19C) y G29645T (V30L).

5. CONCLUSIONES

En el inicio de los primeros casos registrados en España desde finales de enero de 2020 hasta finales de 2021, el virus ha sufrido distintos cambios que se pueden resumir en estos puntos:

1. Todos los virus son mutantes para el gen *orflab* desde 2020, por lo tanto, es probable que todos los virus fueran originalmente mutantes cuando llegaron a España.
2. La tasa de mutación media aumenta alrededor de 5 veces de 2020 (2.99E-4 mutaciones/posición/año) a 2021 (1.57E-3 mutaciones/posición/año). Los genes *orf10* y *orf8* son los que tienen las tasas de mutación más altas en 2020 y 2021 respectivamente, mientras que los genes *orf7a* y *e* tienen las tasas más bajas. Castilla y León tiene los valores más altos entre 2020 y 2021, mientras que Madrid siempre tiene los más bajos.
3. La concentración de las mutaciones se incrementa desde 2020 hasta 2021. Durante 2021, las mutaciones se dispersan más en el 3^{er} cuatrimestre que en el 1º y 2º.
4. La proporción de mutaciones no sinónimas por cada sinónima se ha mantenido constante con un valor aproximado de 1.5:1 durante 2020 y 2021.
5. Entre las mutaciones más relevantes: las transiciones superan a las transversiones y las delecciones empiezan a ser apreciables en el 1^{er} cuatrimestre de 2021.
6. La selección del genoma vírico (dN/dS) ha pasado de tener selección purificadora en 2020 (0.88) a positiva en el 2021 (1.25).
7. El gen *orf8* destaca por tener valores muy altos en la tasa de mutación, concentración de las mutaciones, razón NS/SY, porcentaje de delecciones relevantes, relación Ti/Tv y selección del gen (dN/dS).
8. Los genes *orf6*, *e* y *m* se caracterizan por tener bajas tasas de mutación y valores de selección (dN/dS) bajos que indican selección purificadora.
9. Las mutaciones relevantes más comunes entre todos los mutantes de 2020 y 2021 son C3037T (F924F), C14408T (P4715L) y A23403G (D614G).
10. No existen diferencias significativas entre los genes estructurales y no estructurales en los valores de la tasa de mutación, concentración de las mutaciones, razón NS/SY, razón Ti/Tv y selección genética (dN/dS) tanto en 2020 como en 2021.

6. MATERIAL SUPLEMENTARIO

Ver **ANEXO I, II y III**.

7. REFERENCIAS

- Baric R. S. (2020). *Emergence of a Highly Fit SARS-CoV-2 Variant*. The New England journal of medicine, 383(27), 2684–2686. <https://doi.org/10.1056/NEJMcibr2032888>
- Benedetti, F., Snyder, G.A., Giovanetti, M. et al. (2020). *Emerging of a SARS-CoV-2 viral strain with a deletion in nsp1*. J Transl Med 18, 329. <https://doi.org/10.1186/s12967-020-02507-5>
- Bianchi, M., Borsetti, A., Ciccozzi, M. y Pasarella, S. (2021). *SARS-CoV-2 ORF3a: Mutability and function*. International journal of biological macromolecules, 170, 820–826. <https://doi.org/10.1016/j.ijbiomac.2020.12.142>
- Boson, B., Legros, V., Zhou, B., Siret, E. et al. (2021). *The SARS-CoV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles*. The Journal of biological chemistry, 296, 100111. <https://doi.org/10.1074/jbc.RA120.016175>
- Braun, K. M., Moreno, G. K., Halfmann, P. J., Hodcroft, E. B. et al. (2021). *Transmission of SARS-CoV-2 in domestic cats imposes a narrow bottleneck*. bioRxiv: the preprint server for biology, 2020.11.16.384917. <https://doi.org/10.1101/2020.11.16.384917>
- Cao, Y., Yisimayi, A., Jian, F., Song, W. et al. (2022). *BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection*. Nature. <https://doi.org/10.1038/s41586-022-04980-y>
- Centro de Coordinación de Alertas y Emergencias Sanitarias. (2021). *Variantes de SARS-CoV-2 en España: Omicron*. 8^a actualización.
- Chan, J. F., Yuan, S., Kok, K. H., To, K. K. et al. (2020). *A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster*. Lancet (London, England), 395(10223), 514–523. [https://doi.org/10.1016/S0140-6736\(20\)30154-9](https://doi.org/10.1016/S0140-6736(20)30154-9)
- Chen, H., Mao, Y., Duan, Z., Wang, L. et al. (2021). *Three Cases of COVID-19 Variant Delta With and Without Vaccination*. Chengdu City, Sichuan Province, China CDC weekly, 3(25), 544–546. <https://doi.org/10.46234/ccdcw2021.137>
- Chu, D., Hui, K., Gu, H., Ko, R. et al. (2021). *Introduction of ORF3a-Q57H SARS-CoV-2 Variant Causing Fourth Epidemic Wave of COVID-19, Hong Kong, China*. Emerg. Infect Dis. 2021;27(5):1492-1495. <https://doi.org/10.3201/eid2705.210015>
- Coccia, M. (2022). *Meta-analysis to explain unknown causes of the origins of SARS-CoV-2*. Environmental Research. 211. <https://doi.org/10.1016/j.envres.2022.113062>
- Daly, J. L., Simonetti, B., Klein, K., Chen, K. et al. 2020. *Neuropilin-1 is a host factor for SARS-CoV-2 infection*. Science. 370, 6518: 861-865. <https://doi.org/10.1126/science.abd3072>
- Demir, A. B., Benvenuto, D., Abacioğlu, H., Angeletti, S. y Ciccozzi, M. (2020). *Identification of the nucleotide substitutions in 62 SARS-CoV-2 sequences from Turkey*. Turkish journal of biology = Turk biyoloji dergisi, 44(3), 178–184. <https://doi.org/10.3906/biy-2005-69>
- Di Giacomo, S., Mercatelli, D., Rakhimov, A. y Giorgi, F. (2021). *Preliminary report on SARS-CoV-2 Spike mutation T478K*. bioRxiv. <https://doi.org/10.1101/2021.03.28.437369>
- Desingu, P.A. y Nagarajan, K. (2022), *Omicron BA.2 lineage spreads in clusters and is concentrated in Denmark*. J Med Virol, 94, 2360-2364. <https://doi.org/10.1002/jmv.27659>

Dussupt, V., Sankhala, R.S., Mendez-Rivera, L., Townsley, S.M. et al. (2021). *Low-dose in vivo protection and neutralization across SARS-CoV-2 variants by monoclonal antibody combinations*. Nat Immunol 22. 1503–1514. <https://doi.org/10.1038/s41590-021-01068-z>

Fang, S., Liu, S., Shen, J., Lu, A. et al. (2021). *Updated SARS-CoV-2 Single Nucleotide Variants and Mortality Association*. Med Virol, 93, 6525- 6534. <https://doi.org/10.1002/jmv.27191>

Gao, Y., Yan, L., Huang, Y., Liu, F. et al. (2020). *Structure of the RNA-dependent RNA polymerase from COVID-19 virus*. Science, 368(6492), 779–782. <https://doi.org/10.1126/science.abb7498>

Gao, X., Zhu, K., Qin, B., Olieric, V. et al. (2021). *Crystal structure of SARS-CoV-2 Orf9b in complex with human TOM70 suggests unusual virus-host interactions*. Nat Commun 12, 2843. <https://doi.org/10.1038/s41467-021-23118-8>

Ghosh, S., Dellibovi-Ragheb, T. A., Kerviel, A., Pak, E. et al. (2020). *β-Coronaviruses Use Lysosomes for Egress Instead of the Biosynthetic Secretory Pathway*. Cell 183(6), 1520–1535.e14. <https://doi.org/10.1016/j.cell.2020.10.039>

Ginex, T., Marco-Marín, C., Wieczór, M., Mata, C. et al. (2021). *The structural role of SARS-CoV-2 genetic background in the emergence and success of spike mutations: the case of the spike A222V mutation*. bioRxiv. <https://doi.org/10.1101/2021.12.05.471263>

GISAID (2022). Disponible en: <https://www.gisaid.org/> (Accedido desde el 10 de diciembre de 2020)

Gordon, D.E., Jang, G.M., Bouhaddou, M., Obernier, K. et al. (2020). *A SARS-CoV-2 protein interaction map reveals targets for drug repurposing*. Nature 583, 459–468. <https://doi.org/10.1038/s41586-020-2286-9>

Gordon, C. J., Tchesnokov, E. P., Woolner, E., Perry, J. K. et al. (2020). *Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency*. The Journal of biological chemistry, 295(20), 6785–6797. <https://doi.org/10.1074/jbc.RA120.013679>

Haddad, D., John, S.E., Mohammad, A., Hammad, M.M. et al. (2021) *SARS-CoV-2: Possible recombination and emergence of potentially more virulent strains*. PLoS ONE 16(5). <https://doi.org/10.1371/journal.pone.0251368>

Hammer, Ø., Harper, D. A. T. y Ryan, P. D. (2001). *PAST: Paleontological Statistics Software Package for Education and Data Analysis*. Palaeontologia Electronica, 4, 9. Disponible en: http://palaeo-electronica.org/2001_1/past/issue1_01.htm (Accedido: 10 de junio de 2021).

Hassan, S. S., Aljabali, A., Panda, P. K., Ghosh, S. et al. (2021). *A unique view of SARS-CoV-2 through the lens of ORF8 protein*. Computers in biology and medicine, 133. <https://doi.org/10.1016/j.combiomed.2021.104380>

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N. et al. (2020). *SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor*. Cell, 181(2), 271–280. <https://doi.org/10.1016/j.cell.2020.02.052>

Hossain, M. S., Pathan, A., Islam, M. N., Tonmoy, M. et al. (2021). *Genome-wide identification and prediction of SARS-CoV-2 mutations show an abundance of variants: Integrated study of bioinformatics and deep neural learning*. Informatics in medicine unlocked, 27. <https://doi.org/10.1016/j.imu.2021.100798>

Jagtapa, S., Ratnasri, K., Valloly, P., Sharma, R. et al. (2021). *Evaluation of spike protein antigens for SARS-CoV-2 serology*. Journal of Virological Methods. 296. <https://doi.org/10.1016/j.jviromet.2021.114222>

Jörrißen, P., Schütz, P., Weiand, M., Vollenberg, R. et al. (2021). *Antibody Response to SARS-CoV-2 Membrane Protein in Patients of the Acute and Convalescent Phase of COVID-19*. Frontiers in immunology, 12. <https://doi.org/10.3389/fimmu.2021.679841>

Kosakovsky, S. L. y Frost, S. (2005). *Datammonkey: rapid detection of selective pressure on individual sites of codon alignments*. Bioinformatics, 21(10), 2531–2533. <https://doi.org/10.1093/bioinformatics/bti320>

Kailing, L., Audrey, W., Sheng, L., Shuyi, F. et al. (2022). *Advanced Functions Embedded in the Second Version of Database, Global Evaluation of SARS-CoV-2/hCoV-19 Sequences*. Frontiers in Medicine, 9. ISSN=2296-858X. <https://doi.org/10.3389/fmed.2022.813964>

Kelly, J. A., Olson, A. N., Neupane, K., Munshi, S. et al. (2020). *Structural and functional conservation of the programmed -1 ribosomal frameshift signal of SARS coronavirus 2 (SARS-CoV-2)*. The Journal of biological chemistry, 295(31), 10741–10748. <https://doi.org/10.1074/jbc.AC120.013449>

Kryazhimskiy, S. y Plotkin, J. B. (2008). *The population genetics of dN/dS*. PLoS genetics, 4(12). <https://doi.org/10.1371/journal.pgen.1000304>

Konno, Y., Kimura, I., Uriu, K., Fukushi, M. et al. (2020). *SARS-CoV-2 ORF3b Is a Potent Interferon Antagonist Whose Activity Is Increased by a Naturally Occurring Elongation Variant*. Cell reports, 32(12). <https://doi.org/10.1016/j.celrep.2020.108185>

Kumar, S., Stecher, G. y Tamura, K. (2020). *MEGA11: Molecular Evolutionary Genetics Analysis (version 11.0.10) [Programa de ordenador]*. Molecular Biology and Evolution. 38: 3022-3027. Disponible en: https://www.megasoftware.net/download_form (Accedido: 26 de noviembre de 2021).

Lan, J., Ge, J., Yu, J., Shan, S. et al. (2020). *Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor*. Nature, 581(7807), 215–220. <https://doi.org/10.1038/s41586-020-2180-5>

Leary, S., Gaudieri, S., Parker, M., Chopra, A. et al. (2020). *Generation of a novel SARS-CoV-2 sub-genomic RNA due to the R203K/G204R variant in nucleocapsid: homologous recombination has potential to change SARS-CoV-2 at both protein and RNA level*. bioRxiv. <https://doi.org/10.1101/2020.04.10.209454>

Lee, B., Saqib Sohail, M., Finney, E., Ahmed, S. et al. (2021). *Inferring effects of mutations on SARS-CoV-2 transmission from genomic surveillance data*. MedRxiv. <https://doi.org/10.1101/2021.12.31.21268591>

Shen, L., Triche, T.J., Bard, J.D., Biegel, J.A. et al. (2021). *Spike Protein NTD mutation G142D in SARS-CoV-2 Delta VOC lineages is associated with frequent back mutations, increased viral loads, and immune evasion*. medRxiv. <https://doi.org/10.1101/2021.09.12.21263475>

Lim, I. y Gautheret-Dejean, A. (2021). *Tests rapides antigéniques SARS-CoV-2: définition, législation d'utilisation, principes technologiques, comparaison des performances analytiques et cliniques [Rapid SARS-CoV-2 antigenic test: definition, legislation of use, technological principles, analytical and clinical performance comparison]*. Annales de biologie clinique, 79(2), 123–142. <https://doi.org/10.1684/abc.2021.1635>

Lippi, G., Lavie, C., Henry, B. y Sanchis-Gomar, F. (2020). *Do genetic polymorphisms in angiotensin converting enzyme 2 (ACE2) gene play a role in coronavirus disease 2019 (COVID-19)?* Clinical Chemistry and Laboratory Medicine (CCLM), 58(9), 1415-1422. <https://doi.org/10.1515/cclm-2020-0727>

Lyons, D. M. y Lauring, A. S. (2017). *Evidence for the Selective Basis of Transition-to-Transversion Substitution Bias in Two RNA Viruses*. Molecular biology and evolution, 34(12), 3205–3215. <https://doi.org/10.1093/molbev/msx251>

Malone, B., Urakova, N., Snijder, E.J. y Campbell, E.A. (2022). *Structures and functions of coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug design*. Nat Rev Mol Cell Biol, 23, 21–39. <https://doi.org/10.1038/s41580-021-00432-z>

Miao, G., Zhao, H., Li, Y., Ji, M. et al. (2021). *ORF3a of the COVID-19 virus SARS-CoV-2 blocks HOPS complex-mediated assembly of the SNARE complex required for autolysosome formation*. Developmental cell, 56(4), 427–442. <https://doi.org/10.1016/j.devcel.2020.12.010>

Ministerio de sanidad, Gobierno de España. (2021). *GIV COVID-19: Gestión Integral de la vacunación COVID-19*. Informe para comunicación. Disponible en: https://www.sanidad.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/documentos/Informe_GIV_comunicacion_20210112.pdf (Accedido: 4 de abril de 2021)

Ministerio de Sanidad. (2022). *Variantes de SARS-CoV-2 en España: linajes BA.2.12.1, BA.4 y BA.5 de ómicron. Evaluación rápida de riesgo.* 11^a actualización. Disponible en: <https://www.sanidad.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/documentos/20220628-ERR.pdf> (Accedido: 2 de julio de 2022)

Miorin, L., Kehrer, T., Sanchez-Aparicio, M. T., Zhang, K. et al. (2020). *SARS-CoV-2 Orf6 hijacks Nup98 to block STAT nuclear import and antagonize interferon signaling*. Proceedings of the National Academy of Sciences of the United States of America, 117(45), 28344–28354. <https://doi.org/10.1073/pnas.2016650117>

Mishra, D.S. (2020). *ORF10: Molecular Insights into the Contagious Nature of Pandemic Novel Coronavirus 2019-nCoV*. ChemRxiv. <https://doi.org/10.26434/chemrxiv.12118839.v3>

Miyata, T., Yasunaga, T. y Nishida, T. (1980). *Nucleotide sequence divergence and functional constraint in mRNA evolution*. PNAS, 77(12), 7328-7332. <https://doi.org/10.1073/pnas.77.12.7328>

Mohapatra, R.K., Kandi, V., Tuli, H.S., Chakraborty, C. y Dhama, K. (2022). *The recombinant variants of SARS-CoV-2: Concerns continues amid COVID-19 pandemic*. J Med Virol, 94: 3506-3508. <https://doi.org/10.1002/jmv.27780>

Mou, K., Abdalla, M., Qing, D., Tahir, M. et al. (2021). *Emerging mutations in envelope protein of SARS-CoV-2 and their effect on thermodynamic properties*. Informatics in Medicine Unlocked, 25, ISSN 2352-9148. <https://doi.org/10.1016/j.imu.2021.100675>

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T. et al. (2012). *Detecting Individual Sites Subject to Episodic Diversifying Selection*. PLOS Genetics, 8(7). <https://doi.org/10.1371/journal.pgen.1002764>. Disponible en: <https://www.datammonkey.org/meme> (Accedido: 8 de febrero de 2022).

Nassar, M., Malluhi, Q., Atallah, M. y Shikfa, A. (2017). *Securing Aggregate Queries for DNA Databases*. IEEE Transactions on Cloud Computing, PP. 1-1. <https://doi.org/10.1109/TCC.2017.2682860>

Neches, R. Y., Kyripides, N. C. y Ouzounis, C. A. (2021). *Atypical Divergence of SARS-CoV-2 orf8 from orf7a within the Coronavirus Lineage Suggests Potential Stealthy Viral Strategies in Immune Evasion*. mBio 12(1). <https://doi.org/10.1128/mBio.03014-20>

Nei, M. y Gojobori, T. (1986). *Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions*. Molecular Biology and Evolution, 3(5), 418–426. <https://doi.org/10.1093/oxfordjournals.molbev.a040410>

Nieto-Torres, J. L., DeDiego, M. L., Verdiá-Báguena, C., Jimenez-Guardeño, J. et al. (2014). *Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis*. PLoS pathogens, 10(5). <https://doi.org/10.1371/journal.ppat.1004077>

Nieto-Torres, J. L., Verdiá-Báguena, C., Jimenez-Guardeño, J. M., Regla-Nava, J. A. et al. (2015). *Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome*. Virology, 485, 330–339. <https://doi.org/10.1016/j.virol.2015.08.010>

Okonechnikov, K., Golosova, O. y Fursov, M. (2021). *Unipro UGENE: a unified bioinformatics toolkit (version 41) [Programa de ordenador]*. Bioinformatics, 2012 28: 1166-1167. <https://doi.org/10.1093/bioinformatics/bts091>. Disponible en: <http://ugene.net/download-all.html> (Accedido: 8 de febrero de 2022).

Pancer, K., Milewska, A., Owczarek, K., Dabrowska, A. et al. (2020). *The SARS-CoV-2 ORF10 is not essential in vitro or in vivo in humans*. PLoS pathogens, 16(12). <https://doi.org/10.1371/journal.ppat.1008959>

Patchsung, M., Homchan, A., Aphicho, K., Suraritdechachai, S. et al. (2022). *A multiplexed Cas13-based assay with point-of-care attributes for simultaneous COVID-19 diagnosis and variant surveillance*. medRxiv. <https://doi.org/10.1101/2022.03.17.22272589>

Peacock, T.P., Goldhill, D.H., Zhou, J., Baillon, L. et al. (2021). *The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets*. Nat. Microbiol. 6. <https://doi.org/10.1038/s41564-021-00908-w>

Planas, D., Veyer, D., Baidaliuk, A., Staropoli, I. et al. (2021). *Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization*. Nature, 596, 276–280. <https://doi.org/10.1038/s41586-021-03777-9>

Premkumar, L., Segovia-Chumbez, B., Jadi, R., Martinez, D. R. et al. (2020). *The receptor binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients*. Science immunology, 5(48). <https://doi.org/10.1126/sciimmunol.abc8413>

Public Health England. (2020). *Technical briefing 2. Investigation of novel SARS-CoV-2 variant*. Disponible en: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/959361/Technical Briefing VOC202012-2 Briefing 2.pdf (Accedido: 20 de enero de 2021)

Public Health England. (2021). *Technical Briefing 9. Investigation of novel SARS-CoV-2 variant*. Disponible en: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/979818/Variants_of_Concern_VOC_Technical_Briefing_9_England.pdf (Accedido: 8 de junio de 2021)

Rehman, S., Mahmood, T., Aziz, E. y Batool, R. (2020). *Identification of novel mutations in SARS-CoV-2 isolates from Turkey*. Arch Virol 165, 2937–2944. <https://doi.org/10.1007/s00705-020-04830-0>

Redondo, L., Zaldívar-López, S., Garrido, J. y Montoya M. (2021). *SARS-CoV-2 Accessory Proteins in Viral Pathogenesis: Knowns and Unknowns*. Front. Immunol. <https://doi.org/10.3389/fimmu.2021.708264>

Ren, Y., Shu, T., Wu, D., Mu, J. et al. (2020). *The ORF3a protein of SARS-CoV-2 induces apoptosis in cells*. Cell Mol Immunol, 17, 881–883. <https://doi.org/10.1038/s41423-020-0485-9>

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S. et al. (2017). *DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets*. Molecular biology and evolution, 34(12), 3299–3302. <https://doi.org/10.1093/molbev/msx248>. Disponible en: <http://www.ub.edu/dnasp/> (Accedido: 26 de noviembre de 2021).

Sallard, E., Halloy, J., Casane, D., Decroly, E. et al. (2021). *Tracing the origins of SARS-CoV-2 in coronavirus phylogenies: a review*. Environ Chem Lett. 19, 769–785. <https://doi.org/10.1007/s10311-020-01151-1>

Sarkar, R., Mitra, S., Chandra, P., Saha, P. et al. (2021). *Comprehensive analysis of genomic diversity of SARS-CoV-2 in different geographic regions of India: an endeavour to classify Indian SARS-CoV-2 strains on the basis of co-existing mutations*. Archives of virology, 166(3), 801–812. <https://doi.org/10.1007/s00705-020-04911-0>

Sievers, F. y Higgins, D. G. (2018). *Clustal Omega for making accurate alignments of many protein sequences*. Protein science: a publication of the Protein Society, 27(1), 135–145. <https://doi.org/10.1002/pro.3290>

Sofia, A., Oliveira, F., Shoemark, D. K., Davidson, A.D. et al. (2022). *SARS-CoV-2 spike variants differ in their allosteric response to linoleic acid*. bioRxiv. <https://doi.org/10.1101/2022.04.21.489022>

Stertz, S., Reichelt, M., Spiegel, M., Kuri, T. et al. (2007). *The intracellular sites of early replication and budding of SARS-coronavirus*. Virology. 361(2), 304–315. <https://doi.org/10.1016/j.virol.2006.11.027>

Szemiel, A.M., Merits, A., Orton, R.J., MacLean, O.A. et al. (2021). *In vitro selection of Remdesivir resistance suggests evolutionary predictability of SARS-CoV-2*. PLOS Pathogens, 17(9). <https://doi.org/10.1371/journal.ppat.1009929>

Szymczak, A., Jędruchniewicz, N., Torelli, A., Kaczmarzyk-Radka et al. (2021). *Antibodies specific to SARS-CoV-2 proteins N, S and E in COVID-19 patients in the normal population and in historical samples*. The Journal of general virology. 102 (11), 001692. <https://doi.org/10.1099/jgv.0.001692>

Taylor, J. K., Coleman, C. M., Postel, S., Sisk, J. M. et al. (2015). *Severe Acute Respiratory Syndrome Coronavirus ORF7a Inhibits Bone Marrow Stromal Antigen 2 Virion Tethering through a Novel Mechanism of Glycosylation Interference*. Journal of virology, 89(23), 11820–11833. <https://doi.org/10.1128/JVI.02274-15>

The Perl Foundation. (2021) *The Perl programming language. (version 5.32.1) [Programa de ordenador]*. Disponible en: <https://www.perl.org/get.html> (Accedido: 23 de junio de 2021).

Tchesnokova, V., Kulakesara, H., Larson, L., Bowers, V. et al. (2021). *Acquisition of the L452R mutation in the ACE2-binding interface of Spike protein triggers recent massive expansion of SARS-CoV-2 variants*. bioRxiv. <https://doi.org/10.1101/2021.02.22.432189>

Teoh, K., Siu, Y., Chan, W., Schlüter, M. et al. (2010). *The SARS Coronavirus E Protein Interacts with PALS1 and Alters Tight Junction Formation and Epithelial Morphogenesis*. Molecular Biology of the Cell. 21(22), 3838–3852. <https://doi.org/10.1091/mbc.e10-04-0338>

Vijgen, L., Keyaerts, E., Moës, E., Thoelen, I. et al. (2005). *Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event*. Journal of virology. 79(3), 1595–1604. <https://doi.org/10.1128/JVI.79.3.1595-1604.2005>

ViralZone (2020). Disponible en: <https://viralzone.expasy.org/9076> (Accedido el 16 de abril de 2022)

Wassenaar, T.M., Wanchai, V., Buzard, G. y Ussery, D.W. (2022). *The first three waves of the Covid-19 pandemic hint at a limited genetic repertoire for SARS-CoV-2*. FEMS Microbiology Reviews. <https://doi.org/10.1093/femsre/fuac003>

Wrobel, A.G., Benton, D.J., Xu, P. et al. (2020). *SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects*. Nat Struct Mol Biol, 27, 763–767. <https://doi.org/10.1038/s41594-020-0468-7>

Xi, B., Jiang, D., Li, S., Lon, J. R. et al. (2021). *AutoVEM: An automated tool to real-time monitor epidemic trends and key mutations in SARS-CoV-2 evolution*. Computational and structural biotechnology journal, 19, 1976–1985. <https://doi.org/10.1016/j.csbj.2021.04.002>

Xia, H., Cao, Z., Xie, X., Zhang, X. et al. (2020). *Evasion of Type I Interferon by SARS-CoV-2*. Cell reports, 33(1). <https://doi.org/10.1016/j.celrep.2020.108234>

Yang, R., Zhao, Q., Rao, J., Zeng, F. et al. (2021). *SARS-CoV-2 Accessory Protein ORF7b Mediates Tumor Necrosis Factor-α-Induced Apoptosis in Cells*. Frontiers in microbiology, 12. <https://doi.org/10.3389/fmicb.2021.654709>

Zhao, L. P., Lybrand, T. P., Gilbert, P. B., Hawn, T. R. et al. (2021). *Tracking SARS-CoV-2 Spike Protein Mutations in the United States (2020/01 - 2021/03) Using a Statistical Learning Strategy*. bioRxiv. <https://doi.org/10.1101/2021.06.15.448495>

Zhao, L.P., Roychoudhury, P., Gilbert, P., Shiffer, J. et al. (2022). *Mutations in viral nucleocapsid protein and endoRNase are discovered to associate with COVID19 hospitalization risk*. Sci Rep. 12, 1206. <https://doi.org/10.1038/s41598-021-04376-4>

Zinzula, L. (2021). Lost in deletion: The enigmatic ORF8 protein of SARS-CoV-2. Biochemical and biophysical research communications, 538, 116-124. <https://doi.org/10.1016/j.bbrc.2020.10.045>

ANEXO I – Mutaciones relevantes de los genes de SARS-CoV-2 en Castilla y León, Madrid, Andalucía y Cataluña entre 2020 y los 3 cuatrimestres de 2021 con sus respectivos porcentajes en el intervalo temporal analizado.

orfLab	
	2020
CASTILLA Y LEÓN: C366T (0,4%), T445C (9%), C854T (0,6%), A1547G (2,2%), A1988G (0,7%), C2005T (1,5%), C3037T (14%), G3109T (0,5%), G4006T (0,5%), C5170T (0,5%), C5388A (14,3%), C5700T (4,3%), C6286T (9%), T6294C (4,1%), C7113T (0,6%), T7236C (0,5%), C8782T (0,7%), T9477A (0,6%), C9891T (0,6%), G9929A (1,1%), G11132T (0,5%), C11151T (4,2%), C12415T (0,5%), C14408T (14%), C14805T (0,6%), A15325G (1,1%), G16636T (0,5%), T18471C (1,3%), A20268G (4,6%), G21255C (7,4%). MADRID: T445C (2,8%), C3037T (26,7%), C4539T (0,7%), C6286T (2,9%), C8782T (8,7%), T9477A (2,1%), G11083T (1,7%), T12364A (0,6%), C12970T (0,6%), C14408T (26%), C14805T (2,6%), A19719T (0,6%), A20268G (21,5%), G21255C (2,6%). ANDALUCÍA: T445C (9,9%), G1820A (0,5%), A1988G (0,6%), C3037T (17%), C3096T (0,6%), C3425T (0,6%), C4543T (0,8%), C5170T (0,8%), C5192T (0,6%), G5629T (0,8%), C5784T (0,7%), C6286T (9,8%), C8782T (3,3%), T9477A (2,7%), G9526T (0,8%), C10833T (4,5%), G11083T (0,9%), G11132T (0,8%), C11497T (0,8%), G13993T (0,8%), C14408T (17%), C14805T (2,9%), G15766T (0,8%), C16308T (4,2%), A16889G (0,8%), G17019T (0,8%), G18379T (0,7%), C18877T (0,8%), A20268G (4,5%), T20661C (0,9%), G21255C (9,6%). CATALUÑA: C313T (1,7%), C1059T (2,3%), C3037T (33,8%), G5230T (4,8%), C14408T (32,7%), A20268G (24,8%).	
	2021-Q1
CASTILLA Y LEÓN: T445C (4,5%), C913T (3,2%), C2453T (1,5%), C3037T (8,3%), C3267T (3,2%), A3871G (2,1%), C5388A (5,1%), C5497T (2,2%), C5986T (3,2%), C6286T (4,6%), T6954C (3,2%), C8326T (0,8%), T11288- (3,2%), C11289- (3,2%), T11290- (3,2%), G11291- (3,2%), T11293- (3,2%), T11294- (3,2%), T11296- (3,2%), C11747T (1,5%), C12415T (2,2%), C14408T (8,3%), C14676T (3,2%), C15279T (3,2%), T16176C (3,2%), A17013G (2,1%), C19488T (1,3%), G21255C (4,4%). MADRID: T445C (3,3%), C913T (3,8%), C1059T (0,9%), A1180G (0,7%), C2453T (0,4%), C3037T (8,1%), C3267T (3,7%), T3803C (0,4%), C5170T (0,5%), C5388A (3,7%), C5986T (3,7%), C6286T (3,3%), C6636T (0,4%), C6730T (0,7%), T6954C (4,4%), C6997T (0,3%), G7743T (0,9%), C8152T (0,3%), T8603C (0,5%), C10029T (0,8%), C10207T (0,9%), G10870T (0,4%), G11083T (0,4%), G11132T (0,5%), T11288- (3%), C11289- (3%), T11290- (3%), G11291- (3%), G11292- (3%), T11293- (3%), T11294- (3%), T11295- (3%), T11296- (3%), C11514T (0,7%), C12400T (0,7%), C12439T (0,9%), C13515T (0,9%), C14408T (8,1%), C14676T (3,7%), C15279T (3,7%), T16176C (3,7%), C17172T (0,4%), A17615G (1,5%), G20102T (0,4%), C20187T (0,4%), T20748C (0,8%), G21255C (3,1%). ANDALUCÍA: T445C (2,2%), C913T (3,8%), C1387T (0,4%), C2453T (0,5%), C3037T (6,5%), C3267T (3,8%), C4543T (0,3%), T4844C (1,2%), C5192T (0,5%), C5388A (3,8%), G5629T (0,3%), C5957T (0,3%), C5986T (3,9%), C6286T (2,2%), T6954C (3,9%), A8181G (0,7%), C9073T (0,3%), C9207T (0,3%), G9526T (0,3%), C10833T (1,3%), T11288- (3,8%), C11289- (3,8%), G11291- (3,8%), T11293- (3,8%), T11294- (3,8%), T11295- (3,8%), T11296- (3,8%), C11497T (0,3%), C12100T (0,3%), G12892A (0,5%), C12970T (0,9%), G13993T (0,3%), C14408T (6,5%), C14676T (3,8%), T15096C (1,5%), C15240T (0,3%), C15279T (3,8%), G15766T (0,3%), T16176C (3,7%), C16260T (1,4%), C16308T (1,3%), A16889G (0,3%), G17019T (0,3%), A17615G (0,8%), C18647T (0,5%), C18877T (0,3%), G21255C (2,1%). CATALUÑA: T445C (1%), C913T (4,2%), C1059T (0,3%), G1464A (0,2%), G1820A (0,2%), C2110T (0,7%), C2453T (0,3%), C2544T (0,4%), C3037T (6,1%), C3177T (0,2%), C3267T (4,2%), C4206T (0,2%), C5170T (0,4%), C5301T (0,2%), C5388A (4,2%), A5689G (0,4%), C5986T (4,2%), C6286T (1,1%), C6363T (0,2%), T6954C (4,3%), T7767C (0,2%), C8047T (0,3%), T8156C (0,2%), C10029T (0,2%), A10323G (0,2%), A10978T (0,2%), G11132T (0,4%), G11230A (0,2%), T11288- (4,5%), C11289- (4,5%), T11290- (4,5%), G11291- (4,5%), G11292- (4,5%), T11293- (4,5%), T11294- (4,5%), T11295- (4,5%), T11296- (4,5%), C11750T (0,2%), A12162G (0,3%), C12400T (0,2%), C12809T (0,4%), C14120T (0,6%), C14408T (6,1%), G14547A (0,2%), C14676T (4,2%), G14973A (0,2%), C15279T (4,2%), C15324T (0,4%), T16176C (4,2%), C16726T (0,2%), T17040C (0,2%), C17104T (0,3%), A17615G (1%), A20268G (0,3%), A21137G (0,2%), G21255C (1%).	
	2021-Q2
CASTILLA Y LEÓN: G347A (0,2%), C913T (3,2%), C1191T (0,2%), C1266T (0,2%), G1729A (0,2%), G1738T (0,2%), C2110T (0,5%), G2257A (0,7%), G2284T (0,4%), A2292C (0,1%), G2413A (0,2%), C2453T (0,8%), C2544T (0,3%), C2878T (0,2%), C3037T (5,3%), C3140T (0,2%), C3267T (3,2%), G3638T (0,2%), G4181T (1,7%), C4206T (0,3%), C4543T (0,5%), A4860G (0,1%), C5184T (0,4%), C5388A (2,1%), A5584G (0,2%), A5689G (0,3%), C5986T (3,2%), C6385T (0,1%), C6402T (1,7%), T6619C (0,2%), T6735G (0,2%), T6954C (3,2%), C7124T (1,6%), G7798T (0,2%), C8986T (1,7%), G9053T (1,7%), G9220A (0,1%), C9891T (0,4%), C10029T (1,7%), C10186T (0,3%), A11201G (1,7%), C11208T (0,2%), T11288- (3,2%), C11289- (3,2%), T11290- (3,2%), G11291- (3,2%), G11292- (3,1%), T11293- (3,1%), T11294- (3,1%), T11295- (3,1%), T11296- (3,2%), A11332G (1,7%), T11418C (0,4%), C11514T (0,2%), C11750T (0,3%), C12073T (0,7%), C12809T (0,3%), T12946C (0,2%), C12970T (0,1%), C13019T (0,2%), G14030A (0,2%), C14120T (0,5%), A14213G (0,2%), C14408T (5,3%), G14547A (0,6%), C14676T (3,2%), T15096C (0,2%), C15237T (0,7%), C15279T (3,2%), C15324T (0,3%), G15451A (2%), C15952A (0,6%), T16176C (3,2%), T16242A (0,2%), C16466T (2%), A17615G (0,2%), C17766T (0,2%), C18086T (0,2%), C18176T (0,1%), C18348T (0,1%), C18508T (0,3%), C18555T (0,6%), C18647T (0,3%), C18687T (0,2%), C18744T (0,3%), C18788T (0,2%), C19220T (1,7%), C20031T (0,1%), A20262G (0,2%), C20703T (0,5%). MADRID: C913T (1,1%), C1191T (0,5%), C1266T (0,4%), C1593T (0,6%), G1729A (0,5%), G1738T (0,4%), G2257A (0,4%), A2292C (0,4%), C3037T (7,3%), C3267T (1,1%), A3428G (0,4%), G4181T (4,4%), C4878T (0,4%), C5184T (1,2%), C5388A (1,1%), A5584G (0,7%), C5986T (1,1%), C6037T (0,4%), C6402T (4,4%), C6840T (0,4%), T6954C (0,9%), C7124T (4,4%), C7851T (0,5%), C8986T (4,3%), G9053T (4,4%), C9891T (1,3%), C10029T (4,8%), A10323G (0,9%), A11201G (4,4%), T11288A (0,8%), C11289A (0,8%), T11290G (0,8%), G11291A (0,8%), T11296G (0,6%), A11332G (4,4%), G11371C (0,8%), T11418C (1,1%), A11451G (0,4%), C11514T (0,8%), C12073T (0,5%), T12946C (0,4%), C13019T (0,6%), A13057T (0,4%), G14030A (0,5%), C14408T (7,1%), C14676T (1,1%), C15237T (0,5%), C15279T (1,1%), G15451A (5,6%), C15952A (1,7%), T16176C (1,1%), C16466T (5,6%), G17193T (0,5%), C17491T (0,4%), C18086T (0,5%), C18176T (0,4%), C18348T (0,4%), C18744T (0,8%), C18877T (0,4%), T19035C (0,3%), C19220T (4,4%), C20031T (0,4%), C20148T (0,3%), A20262G (0,4%), C21301A (0,4%). ANDALUCÍA: C913T (4,3%), C2453T (1,3%), C3037T (5,5%), C3267T (4,3%), G4181T (1%), T4844C (1%), C5388A (4,3%), C5986T (4,3%), C6402T (1%), T6954C (4,3%), C7124T (1%), C8986T (1%), G9053T (1%), C10029T (1%), A11201G (1%), T11288- (4,3%), C11289- (4,3%), T11290- (4,3%), G11291- (4,3%), G11292- (4,3%), T11293- (4,3%), T11294- (4,3%), T11295- (4,3%), T11296- (4,3%), A11332G (1%), G12892A (1%), C14408T (5,5%), C14676T (4,3%), T15096C (1,3%), C15279T (4,3%), G15451A (1,3%), T16176C (4%), C16260T (1%), C16466T (1,3%), C19220T (1%). CATALUÑA: G2257A (1%), C3037T (7,3%), G4181T (6,7%), C6402T (6,7%), C7124T (6,7%), C8986T (6,7%), G9053T (6,7%), C9891T (0,4%), C10029T (6,7%), A11201G (6,7%), A11332G (6,7%), C12073T (1,1%), C14408T (7,3%), C15237T (1,1%), G15451A (7,1%), C15952A (4,4%), C16466T (7,1%), C18744T (2,5%), C19220T (6,7%).	

CASTILLA Y LEÓN: C383T (0,1%), T516- (0,1%), A517- (0,1%), G598A (0,3%), A685- (0,1%), A686- (0,1%), G687- (0,1%), T688- (0,1%), C689- (0,1%), A690- (0,1%), T691- (0,1%), T692- (0,1%), T693- (0,1%), C744T (0,1%), C828T (0,1%), G1047T (0,1%), C1076T (0,4%), C1347T (0,1%), T1371C (0,1%), C1448T (0,1%), C1472T (0,2%), C1479T (0,2%), C1592T (0,2%), C1683T (0,1%), G1729A (0,4%), G1820A (0,1%), C1911T (0,1%), C1918T (0,3%), A2151G (0,1%), G2257A (0,5%), A2832G (0,1%), C3037T (5,8%), C3429T (0,1%), A3569G (0,2%), C3587T (0,2%), G4181T (5,2%), C4321T (0,3%), C5144T (0,1%), C5184T (0,5%), G5230T (0,3%), C5310T (0,1%), T5386G (0,1%), A5584G (0,5%), C5672T (0,1%), C5907T (0,1%), G5924A (0,1%), C6310T (0,1%), G6349T (0,2%), C6402T (5,1%), G6513- (0,1%), T6515- (0,1%), G6532T (0,1%), C6629A (0,3%), C6730T (0,1%), C7119T (0,1%), C7124T (5,1%), G7235T (0,1%), C7851T (0,6%), T7984C (0,1%), C8326T (0,1%), G8393A (0,1%), A8413G (0,2%), C8986T (5,1%), G9053T (5,1%), C9120T (0,1%), C9165T (0,2%), G9441T (0,3%), C9559T (0,1%), T9702C (0,1%), C9891T (0,5%), T9955C (0,1%), C10029T (5,2%), A10323G (0,3%), C10449A (0,1%), C10450T (0,1%), C10977T (0,2%), A11201G (5,1%), T11286- (0,1%), G11287- (0,1%), T11288- (0,1%), C11288- (0,1%), T11289- (0,1%), G11290- (0,1%), A11332G (5,1%), T11418C (0,5%), C11514T (0,6%), A11537G (0,1%), T11560C (0,1%), G11851T (0,3%), C11956T (0,2%), C12073T (0,5%), C12473T (0,2%), A12772G (0,1%), C13019T (0,5%), T13195C (0,1%), A13712G (0,2%), C13944T (0,1%), C13971T (0,2%), G14030A (0,4%), C14117T (0,3%), G14118A (0,1%), C14403T (0,1%), C14408T (5,7%), C14585T (0,1%), G15006T (0,5%), C15237T (0,5%), C15240T (0,3%), T15264C (0,1%), C15324T (0,1%), C15368T (0,1%), G15451A (5,5%), C15656T (0,1%), G15906T (0,1%), C15952A (2,2%), T16056C (0,1%), C16338T (0,1%), G16377A (0,2%), C16466T (5,5%), C16575T (0,5%), C16887T (0,1%), T16993C (0,1%), T17040C (0,2%), C17463T (0,3%), G17679A (0,3%), T17955G (0,2%), C18029T (0,1%), C18086T (0,4%), A18163G (0,1%), G18255T (0,4%), A18298G (0,5%), T18438C (0,1%), C18647T (0,2%), C18687T (0,1%), C18744T (1,6%), G18905A (0,1%), G19072T (0,1%), C19220T (5,1%), G19542T (0,1%), C19545T (0,5%), G19563A (0,3%), G19677T (0,5%), C19718T (0,1%), C19983T (0,2%), G20102T (0,1%), G20269T (0,1%), A20481G (0,1%), G20578T (0,1%), G20931T (0,5%), G20937T (0,5%), C21301A (0,2%), A21391G (0,2%). **MADRID:** G526T (0,2%), C828T (0,2%), G1047T (0,3%), T1186A (0,1%), C1266T (0,1%), G1729A (0,2%), G1738T (0,1%), C1758T (0,1%), C1874T (0,3%), G2257A (0,2%), G2518T (0,2%), T2745C (0,1%), C3037T (6,3%), G4181T (6,2%), C4331T (0,4%), C4754T (0,2%), C5184T (0,2%), G5437T (0,2%), A5584G (0,3%), C6402T (6%), G6531T (0,1%), C7124T (6,2%), A7186T (0,2%), C7428T (0,2%), C7851T (0,6%), T8656C (0,3%), C8986T (6,1%), G9053T (6%), C9165T (0,4%), G9190T (0,1%), G9441T (0,2%), C9891T (0,3%), C10029T (6,3%), C10319T (0,2%), A10323G (0,4%), C11151T (0,1%), A11201G (6,1%), A11332G (6,2%), G11371C (2,6%), T11418C (0,2%), C11514T (0,2%), C11747T (0,1%), A11996T (0,1%), C12073T (0,2%), C12439T (0,2%), A12772G (0,3%), C13019T (0,3%), G14030A (0,2%), C14408T (6%), C14793T (0,1%), C15237T (0,2%), C15240T (0,5%), G15451A (6,4%), G15910T (0,1%), C15952A (3%), C16466T (6,3%), T17040C (0,3%), G17259T (0,1%), C17977T (0,1%), C18086T (0,2%), G18255T (0,1%), C18744T (1,9%), C18877T (0,1%), C19220T (6,1%), T19458C (0,1%), C19524T (0,2%), G19563A (0,1%), C20148T (0,1%), C20178T (0,3%), C20233T (0,1%), G20937T (0,1%). **CATALUÑA:** C334T (0,1%), G509- (0,01%), T510- (0,01%), C511- (0,01%), A512- (0,01%), T513- (0,01%), G514- (0,1%), T515- (0,1%), T516- (0,1%), A517- (0,1%), T518- (0,1%), G519- (0,1%), G526T (0,1%), A534T (0,3%), C593T (0,1%), A685- (0,1%), A686- (0,1%), G687- (0,1%), T688- (0,1%), C689- (0,1%), A690T (0,1%), T691- (0,1%), T692- (0,1%), T693- (0,1%), C744T (0,2%), G777A (0,01%), A828G (0,3%), C828T (0,2%), C913T (0,1%), C933T (0,1%), C1043T (0,1%), G1047T (0,4%), G1124A (0,01%), C1384T (0,1%), C1419T (0,4%), C1472T (0,1%), C1592T (0,01%), C1615T (0,2%), A1690G (0,3%), G1729A (0,1%), G1738T (0,01%), C1758T (0,1%), G1820A (0,1%), C1911T (0,1%), G2019A (0,01%), C2113T (0,1%), T2152C (0,2%), G2246A (0,01%), G2257A (0,1%), A2317G (0,1%), C2395T (0,1%), G2518T (0,1%), T2745C (0,1%), C2788T (0,1%), C2939T (0,1%), C2980T (0,1%), C3037T (5,6%), G3122T (0,01%), G3250T (0,01%), G3340T (0,01%), C3738T (0,01%), C3773T (0,3%), G3870T (0,1%), C3874T (0,1%), G3955T (0,1%), C4113T (0,01%), G4181T (5,4%), G4201T (0,01%), A4269C (0,01%), C4276T (0,01%), G4309T (0,01%), C4331T (0,01%), C4539T (0,1%), C4543T (0,01%), C4683T (0,01%), C4780A (0,1%), A4783T (0,1%), C4965T (0,01%), G4975A (0,01%), C5100T (0,2%), C5184T (0,1%), G5326T (0,01%), C5365T (0,1%), C5392T (0,1%), G5437T (0,1%), C5497T (0,01%), A5584G (0,1%), C5826T (0,1%), C5835T (0,01%), C6070T (0,01%), G6080T (0,01%), C6310T (0,1%), C6400T (0,3%), C6402T (5,4%), C6468G (0,01%), C6500T (0,01%), T6547C (0,01%), A6551G (0,01%), T6580C (0,01%), G6622A (0,4%), C6696T (0,01%), C6730T (0,2%), C6855T (0,01%), C7124T (5,4%), C7564T (0,01%), G7675T (0,3%), C7765T (0,1%), C7768T (0,1%), G7798T (0,1%), C7851T (0,8%), G7883A (0,2%), A7949G (0,01%), C8016T (0,01%), C8326T (0,1%), T8656C (0,1%), T8682C (0,1%), C8986T (5,4%), G9053T (5,4%), G9077A (0,01%), T9139C (0,01%), A9162G (0,1%), C9165T (0,1%), G9190T (0,01%), C9223T (0,1%), C9246T (0,1%), C9344T (0,1%), C9365T (0,1%), C9430T (0,01%), G9441T (0,01%), C9521T (0,1%), C9559T (0,1%), G9856T (0,2%), C9891T (0,1%), C10029T (5,4%), C10138T (0,2%), A10181G (0,3%), A10323G (0,2%), C10449T (0,01%), C10537T (0,1%), C10802T (0,1%), C10833T (0,1%), T10942C (0,01%), C10977T (0,1%), C11003T (0,01%), G11083T (0,1%), G11113T (0,1%), A11201G (5,4%), A11332G (5,4%), T11418C (0,1%), C11494T (0,01%), C11514T (0,1%), T11569G (0,01%), C11931T (0,1%), C11941T (0,01%), C12025T (0,2%), C12073T (0,1%), C12484T (0,1%), A12573G (0,01%), G12769A (0,1%), A12772G (0,1%), A12949G (0,1%), C13019T (0,1%), C13168T (0,2%), C13252T (0,01%), C13329T (0,01%), G13459A (0,01%), C13860T (0,01%), A13898G (0,1%), C13971T (0,01%), G14030A (0,1%), C14408T (5,5%), G14443T (0,3%), A14510G (0,3%), G14511C (0,4%), T14544C (0,1%), G14584A (0,01%), C14724T (0,01%), G14854A (0,1%), C14925T (0,1%), C15237T (0,1%), C15240T (0,2%), C15315A (0,01%), G15451A (5,5%), T15628C (0,3%), C15660T (0,1%), C15720T (0,1%), G15906T (0,2%), C15933T (0,01%), C15952A (1,8%), C16111T (0,01%), C16428T (0,3%), C16466T (5,5%), G16710A (0,1%), C16726T (0,2%), C16887T (0,2%), A16998G (0,3%), C17010T (0,1%), T17011C (0,01%), T17040C (0,5%), G17122T (0,1%), C17133T (0,01%), C17172T (0,01%), A17236G (0,1%), C17304T (0,1%), C17676T (0,1%), C17746T (0,3%), C17747A (0,1%), C18029T (0,01%), G18079T (0,3%), C18086T (0,1%), C18138T (0,01%), G18255T (0,4%), C18312A (0,1%), T18437G (0,01%), G18462A (0,01%), T18465C (0,01%), G18583A (0,1%), C18744T (1,5%), A18846G (0,3%), G18905A (0,2%), C19007T (0,01%), G19072T (0,01%), C19164T (0,01%), C19220T (5,4%), C19245T (0,1%), C19374T (0,1%), A19515G (0,01%), C19524T (0,1%), G19563A (0,01%), T19839C (0,01%), C19859T (0,1%), C19862T (0,1%), G19962T (0,1%), C20178T (0,1%), G20274T (0,1%), G20397T (0,2%), T20501C (0,1%), A20532G (0,3%), A20608G (0,1%), G20683A (0,3%), A20755G (0,01%), T20928A (0,1%), G20937T (0,2%), C21058T (0,4%), A21137G (0,1%), G21160A (0,3%), C21179T (0,01%), A21195G (0,01%), T21196C (0,01%), C21301A (0,01%).

s (spike)

2020

CASTILLA Y LEÓN: T21765- (0,9%), A21766- (0,9%), C21767- (0,9%), A21768- (0,9%), T21769- (0,9%), G21770- (0,9%), G21898T (1,9%), C22227T (33,3%), A23063T (0,9%), C23271A (0,9%), A23403G (52%), C23604A (0,9%), C23709T (0,9%), T24506G (0,9%), G24914C (0,9%), T25316C (1,7%). **MADRID:** A22151G (2%), C22227T (9,1%), A23403G (88,9%). **ANDALUCÍA:** C21575T (0,7%), C21855T (0,9%), T21991- (0,9%), T21992- (0,9%), A21993- (0,9%), A22151G (1%), C22227T (29,8%), T22297C (2%), C22432T (1,8%), C22858T (1,1%), G22933A (2%), C22943A (0,9%), G22992A (2,3%), A23403G (52,7%), C23533T (0,7%), C23730T (0,7%), G24812T (0,7%). **CATALUÑA:** A23403G (100%).

2021-Q1

CASTILLA Y LEÓN: T21765- (5,2%), A21766- (5,2%), C21767- (5,2%), A21768- (5,2%), T21769- (5,2%), G21770- (5,2%), G21974C (2,1%), T21991- (4,9%), T21992- (4,9%), A21993- (4,9%), C22227T (7,5%), A23063T (5,2%), C23271A (5,2%), A23403G (13,4%), C23604A (5,2%), C23709T (5,2%), T24506G (5,2%), G24914C (5,2%). **MADRID:** C21614T (0,3%), T21765- (4,6%), C21767- (4,6%), A21768- (4,6%), T21769- (4,6%), G21770- (4,6%), G21974C (5,2%), C22227T (0,3%), A22151G (2%), C22227T (9,1%), A23403G (88,9%). **ANDALUCÍA:** C21575T (0,7%), C21855T (0,9%), T21991- (0,9%), T21992- (0,9%), A21993- (0,9%), A22151G (1%), C22227T (29,8%), T22297C (2%), C22432T (1,8%), C22858T (1,1%), G22933A (2%), C22943A (0,9%), G22992A (2,3%), A23403G (52,7%), C23533T (0,7%), C23730T (0,7%), G24812T (0,7%). **CATALUÑA:** C21575T (0,3%), C21614T (0,4%), C21621A (0,2%), C21638T (0,2%), T21765- (5,7%), A21766- (5,7%), C21767- (5,7%), A21768- (5,7%), T21769- (5,7%), G21770- (5,7%), G21974C (5,9%), C22227T (0,2%), A22151G (2%), C22227T (1,4%), A22812C (0,2%), C22879A (0,3%), G23012A (0,4%), T23042C (0,3%), A23063T (6%), C23271A (5,7%), A23403G (8,3%), T23407C (0,2%), C23525T (0,2%), C23604A (6,2%), C23664T (0,2%), C23709T (5,9%), G24000A (0,2%), T24506G (5,7%), C24642T (0,2%), G24914C (5,7%), G25088T (0,2%).

2021-Q2	
CASTILLA Y LEÓN: C21618G (2,5%), G21641T (0,2%), T21765- (4%), A21766- (4%), C21767- (4%), A21768- (4%), T21769- (4%), G21770- (4%), C21846T (1,3%), G21974C (1%), G21987A (2,4%), T21991- (3,7%), T21992- (3,8%), A21993- (3,8%), T21994C (0,2%), A22029- (2,4%), G22030- (2,4%), T22031- (2,4%), T22032- (2,4%), C22033- (2,4%), A22034- (2,4%), C22227T (0,2%), C22624T (0,3%), C22792T (0,2%), T22917G (2,5%), C22995A (2,5%), A23063T (4,1%), C23271A (4%), C23277T (0,2%), A23403G (6,6%), C23604A (6,6%), C23709T (4%), C23718T (0,2%), C24130T (0,1%), G24410A (2,6%), T24460C (0,1%), T24506G (4%), C24865T (0,3%), G24914C (4%), G24928T (0,1%), C25169T (0,1%). MADRID: C21614T (0,2%), C21618G (6,9%), G21624C (0,3%), G21641T (0,6%), T21765- (0,4%), A21766- (0,4%), C21767- (0,4%), A21768- (0,4%), T21769- (0,4%), G21770- (0,4%), G21786T (0,2%), C21789T (0,3%), G21987A (3,1%), G21987A (3,4%), T21991- (0,8%), T21992- (0,9%), A21993- (1,3%), T21995A (0,4%), A22029- (4,1%), G22030- (4,1%), T22031- (4,1%), T22032- (4,1%), C22033- (4,1%), A22034- (4,1%), C22082T (0,2%), A22206G (0,2%), C22227T (1%), C22314T (0,2%), G22599A (0,5%), C22792T (0,7%), G22813T (0,2%), T22917G (7,4%), C22987T (0,2%), C22995A (7,3%), G23012A (0,7%), A23063T (2%), C23271A (1,5%), T23398C (0,4%), A23403G (9,8%), G23593T (0,2%), C23604G (9,4%), C23709T (1,4%), C23718T (0,2%), G24040T (0,7%), C24130T (0,2%), C24138A (0,2%), G24410A (6,3%), T24506G (1,4%), G24914C (1,5%), G24928T (0,7%). ANDALUCÍA: C21618G (1,6%), T21765- (5,3%), A21766- (5,3%), C21767- (5,3%), A21768- (5,3%), T21769- (5,3%), G21770- (5,3%), T21991- (3,2%), T21992- (3,2%), A21993- (3,2%), C21998T (0,8%), A22029- (1,6%), G22030- (1,6%), T22031- (1,6%), T22032- (1,6%), C22033- (1,6%), A22034- (1,6%), T22917G (1,6%), C22995A (1,6%), A23063T (5,3%), C23271A (5,3%), C23315T (1,2%), A23403G (6,9%), C23604A (6,9%), C23709T (5,3%), G24410A (1,6%), T24506G (5,3%), G24914C (5,3%). CATALUÑA: C21618G (7,8%), C21846T (1,6%), G21987A (2,7%), A22029- (7,8%), G22030- (7,8%), T22031- (7,8%), T22032- (7,8%), A22034- (7,8%), C22227T (0,4%), C22792T (0,2%), T22917G (7,7%), C22995A (7,8%), A23063T (0,2%), C23277T (0,2%), A23403G (8%), C23604G (8%), C23718T (0,2%), C24130T (0,5%), G24410A (7,2%), G24926T (0,2%), G24928T (0,4%).	
	2021-Q3
CASTILLA Y LEÓN: C21618G (6,8%), A21647G (0,3%), C21846T (2,8%), G21987A (6,8%), T21990- (0,2%), T21991- (0,3%), T21992- (0,3%), A21993- (0,3%), T21994- (0,2%), T21995- (0,2%), A22029- (6,8%), G22030- (6,8%), T22031- (6,8%), C22033- (6,8%), A22034- (6,8%), G22201T (0,2%), C22227T (0,7%), C22311T (0,3%), C22458T (0,3%), C22498T (0,6%), C22624T (0,2%), C22792T (0,5%), G22899T (0,2%), T22917G (6,8%), C22938T (0,2%), C22995A (6,9%), G23012C (0,5%), A23013C (0,7%), A23063T (0,2%), G23401T (0,3%). A23403G (7%), C23604G (7%), C23854A (0,1%), C24130T (0,9%), G24410A (6,8%), T24589C (0,1%), G25352T (0,2%). MADRID: C21614T (0,1%), C21618G (6,7%), A21647G (0,1%), C21846T (2,8%), G21987A (7,7%), T21992A (0,1%), A21993C (0,1%), T21995C (0,3%), A21996C (0,2%), C21997T (0,2%), ins21998A (0,1%), ins21999A (0,1%), ins22000C (0,1%), A22029- (5,8%), G22030- (5,8%), T22031- (5,8%), T22032- (5,8%), C22033- (5,8%), A22034- (5,8%), C2208T (0,2%), C22208T (0,2%), C22227T (0,5%), C22311T (0,1%), C22314T (0,2%), C22498T (0,1%), G22599A (0,1%), C22792T (0,3%), C22858T (0,2%), T22917G (8,1%), C22993T (0,1%), C22995A (8,1%), G23012A (0,2%), A23063T (0,2%), T23189G (0,2%), C23202A (0,1%), G23401T (0,2%), A23403G (8,3%), C23525T (0,1%), G23593C (0,3%), C23604G (8,2%), C23683T (0,1%), C23718T (0,2%), C23741T (0,1%), C24130T (0,8%), G24199A (0,4%), G24410A (8%), C24707A (0,1%), G24872T (0,1%), C25006T (0,1%), T25027C (0,5%), G25337T (0,1%). CATALUÑA: C21575T (0,1%), G21604T (0,4%), C21614T (0,3%), C21618G (6,9%), A21647G (0,01%), C21846T (3,5%), C21897T (0,1%), T21973C (0,4%), G21974C (0,1%), G21987A (6,9%), T21995C (0,5%), A22029- (6,9%), G22030- (6,9%), T22031- (6,9%), C22033- (6,9%), A22034- (6,9%), G22104T (0,1%), C22120T (0,1%), C22208T (0,1%), C22227T (0,7%), C22281T (0,1%), T22307C (0,01%), C22311T (0,01%), C22314T (0,1%), C22413T (0,1%), G22468T (0,1%), C22498T (0,1%), C22792T (0,1%), T22795C (0,4%), T22917G (6,9%), C22993T (0,1%), C22995A (6,9%), G23012C (0,01%), C23277T (0,1%), A23403G (6,9%), C23557T (0,2%), G23593T (0,1%), C23604G (6,9%), C23635T (0,01%), C23638T (0,1%), A23667T (0,1%), C23683T (0,1%), C23896T (0,1%), A24014G (0,1%), G24040T (0,1%), G24095T (0,01%), C24130T (0,5%), G24199A (0,1%), G24410A (6,9%), C24559T (0,1%), G24872T (0,3%), G24928T (0,1%), C25066T (0,1%), C25156T (0,1%), G25186A (0,1%), T25233C (0,1%).	

orf3a

2020	
CASTILLA Y LEÓN: C25613T (38,9%), G25979T (61,1%). MADRID: T25496C (4,3%), G25563T (4,3%), G25979T (19,6%), C26088T (58%), G26144T (13,8%). ANDALUCÍA: C25418A (7,2%), C25517G (7,2%), G25563T (16,9%), C25710T (9,2%), C25777T (8,7%), C25886T (5,3%), G25979T (31,4%), A26059G (3,9%), C26088T (5,8%), G26144T (4,3%). CATALUÑA: T25485C (27,3%), G25563T (72,7%).	

2021-Q1

CASTILLA Y LEÓN: G25494T (15,2%), G25563T (21,2%), G25776T (15,2%), G25855T (12,1%), G26155- (12,1%), T26156- (12,1%), T26157- (12,1%). MADRID: C25413T (30,4%), G25437T (10,9%), G25563T (30,4%), C25904T (17,4%), A25905T (10,9%). ANDALUCÍA: A25505G (6,1%), C25517G (21,2%), G25563T (13,6%), C25626T (10,6%), C25710T (13,6%), C25886T (10,6%), G25906C (6,1%), G25996T (6,1%), C26022T (12,1%). CATALUÑA: G25538A (10,5%), G25563T (15,8%), T25570C (7,9%), T25715A (9,2%), C25795T (9,2%), C25847T (9,2%), C25907T (7,9%), A25961G (9,2%), G25997T (13,2%), T26152C (7,9%).	
--	--

2021-Q2

CASTILLA Y LEÓN: C25433T (2,7%), G25437T (2,7%), C25469T (59,1%), C25872T (20,4%), C26060T (14,9%). MADRID: C25469T (80,2%), G25563T (6,3%), G25660A (1,9%), C25844T (1,9%), C25872T (4,7%), C26060T (1,9%), G26167T (1,6%), G26172T (1,6%). ANDALUCÍA: C25469T (66,7%), C25626T (33,3%). CATALUÑA: C25469T (83,8%), G25472T (3,8%), C25876T (12,4%).	
---	--

2021-Q3

CASTILLA Y LEÓN: C25460T (0,9%), C25469T (64,5%), G25471T (1,5%), T25518C (0,2%), T25518G (0,4%), C25549T (0,9%), C25568T (0,5%), C25584T (1,3%), C25626T (0,5%), G25690T (1,6%), ins25702C (2,7%), ins25703C (2,7%), ins25704C (2,7%), C25714T (1,3%), G25720T (1,1%), C25810T (0,7%), G25855T (0,5%), C25872T (5,5%), C26022T (0,5%), G25690A (0,2%), G25690T (1,5%), G26173T (5,8%). MADRID: C25469T (83,2%), G25471T (1,2%), G25563T (1,2%), C25614T (2,7%), ins25702C (1,5%), ins25703C (1,5%), ins25704C (1,5%), G26109A (2,2%), G26158- (1,2%), T26159- (1,2%), T26160- (1,2%), A26161- (1,2%). CATALUÑA: C25413T (5,3%), G25440T (0,5%), C25463T (0,01%), C25463A (0,7%), C25469T (51,2%), G25471T (1,5%), G25489T (0,4%), T25514C (0,4%), T25518C (3,2%), C25521T (2,8%), G25526T (0,5%), G25538T (3,5%), C25568T (1,4%), C25614T (4,2%), C25624T (0,5%), C25667T (1%), G25690T (0,8%), G25699C (0,1%), G25699A (0,3%), ins25702C (0,4%), ins25703C (0,4%), ins25704T (0,01%), ins25704C (0,3%), G25720T (0,5%), C25721T (1%), G25906T (0,1%), G25906C (0,4%), A25762T (0,5%), G25855T (1,5%), A25856G (0,4%), C25872T (0,9%), G25906C (0,5%), G25907T (0,5%), C25916T (0,5%), G25947T (0,01%), G25947C (0,5%), G25996A (0,01%), G25996T (1,6%), C26013T (0,5%), G26020T (0,6%), C26054T (0,01%), C26054A (3,5%), C26060T (2,6%), G26107C (0,6%), G26109A (0,5%), A26170T (3,1%), G26172T (0,4%).	
---	--

orf3b

2020	
CASTILLA Y LEÓN: C25836T (31,3%), G25979T (68,8%). MADRID: G25979T (21,4%), C26088T (63,5%), G26144T (15,1%). ANDALUCÍA: C25777T (12,9%), C25886T (7,9%), G25906C (5%), G25979T (47,5%), G25996T (5%), A26059G (5,8%), C26088T (8,6%), G26144T (7,2%). CATALUÑA: A25765T (11,1%), T25779A (11,1%), G25979T (22,2%), G25996T (11,1%), C26111A (22,2%), G26144T (22,2%).	

2021-Q1

CASTILLA Y LEÓN: G25776T (23,8%), G25855T (19%), G26155- (19%), T26156- (19%), T26157- (19%). MADRID: C25904T (50%), A25905T (31,3%), G25996T (18,8%). ANDALUCÍA: G25785T (9,8%), C25886T (27,5%), G25906C (15,7%), G25996T (15,7%), C26022T (31,4%). CATALUÑA: C25792T (17,8%), C25844T (15,6%), C25904T (15,6%), A25958G (15,6%), G25996T (22,2%), T26149C (13,3%).	
---	--

2021-Q2

CASTILLA Y LEÓN: G25769T (2,7%), G25855T (4,1%), C25872T (45,3%), C26060T (33,1%), G26158- (3,4%), T26159- (3,4%), T26160- (4,7%), A26161- (3,4%). MADRID: C25844T (16,2%), C25872T (40,5%), C26060T (16,2%), G26167T (13,5%), G26172T (13,5%). ANDALUCÍA: G25947T (16,7%), C26013T (16,7%), C26060T (16,7%), T26103A (16,7%), G26104A (16,7%). CATALUÑA: C25872T (90,6%), C26111T (9,4%).	
--	--

2021-Q3

CASTILLA Y LEÓN: G25767T (1,6%), C25810T (3,3%), C25831T (1,6%), G25855T (2,5%), C25872T (24,6%), G25907T (1,6%), G25912T (1,6%), C26022T (2,5%), G26109A (18%), T26149C (1,6%), A26161- (1,6%), A26162- (1,6%), T26163- (1,6%), C26164- (1,6%), C26165- (1,6%), A26166- (1,6%), G26167- (1,6%), T26168- (1,6%), A26169- (1,6%), G26173T (26,2%). **MADRID:** C25872T (10,8%), C26060T (10,8%), G26109A (24,3%), G26158- (13,5%), T26159- (13,5%), T26160- (13,5%), A26161- (13,5%). **CATALUÑA:** C25810T (0,9%), A25817G (0,9%), C25844T (1,6%), G25855T (7,1%), A25856G (2,1%), C25872T (4,6%), G25906T (0,7%), G25906C (1,8%), G25907T (2,5%), C25916T (2,5%), G25947T (0,2%), G25947C (2,3%), G25996A (0,2%), G25996T (8,3%), C26013T (2,3%), G26020T (3,2%), A26050G (1,1%), C26054T (0,2%), C26054A (17,7%), C26060T (12,8%), G26062T (1,1%), C26078T (0,9%), C26082T (0,9%), G26107C (3%), G26109A (2,3%), A26170T (15,4%), G26172T (1,8%), C26208T (1,4%).

e

2020

CASTILLA Y LEÓN: G26268T (100%). **MADRID:** T26301G (100%). **ANDALUCÍA:** G26268T (54,5%), T26289C (27,3%), T26301G (9,1%). **CATALUÑA:** SIN MUTANTES.

2021-Q1;

CASTILLA Y LEÓN: C26270T (50%), C26464T (50%). **MADRID:** C26456T (100%). **ANDALUCÍA:** G26262A (100%). **CATALUÑA:** T26446G (33,3%), C26456T (66,7%).

2021-Q2;

CASTILLA Y LEÓN: C26270T (50%), C26464T (50%). **MADRID:** C26408T (33,3%), C26456T (66,7%). **ANDALUCÍA:** G26428T (100%). **CATALUÑA:** C26408T (100%).

2021-Q3;

CASTILLA Y LEÓN: C26256T (14,3%), C26270T (85,7%). **MADRID:** A26267G (6,7%), C26270T (13,3%), C26340T (6,7%), G26389T (6,7%), C26408T (6,7%), G26428T (20%), G26439T (6,7%), C26444T (20%), C26455T (6,7%), C26464T (6,7%). **CATALUÑA:** G26365T (16,7%), T26409C (23,3%), G26416T (20%), T26442A (30%), C26455T (10%).

m

2020

CASTILLA Y LEÓN: T26609C (4,8%), C26801G (95,2%). **MADRID:** C26681T (14%), C26801G (86%). **ANDALUCÍA:** C26533T (2%), C26681T (2,3%), C26735T (6,3%), C26801G (80,2%), T26876C (6,3%), G27112A (3%). **CATALUÑA:** C26801G (100%).

2021-Q1

CASTILLA Y LEÓN: G26526T (3,8%), C26735T (6,4%), A26748G (5,1%), C26801G (84,6%). **MADRID:** T26767C (18,6%), C26801G (81,4%). **ANDALUCÍA:** C26735T (10,7%), T26744C (2,5%), C26801G (73%), T26876C (10,7%), C26985T (3,1%). **CATALUÑA:** G26526T (8,6%), T26604C (18,3%), C26636T (10,8%), T26767C (8,6%), C26801G (47,3%), C26936T (6,5%).

2021-Q2

CASTILLA Y LEÓN: G26526T (11,1%), T26604C (9,9%), T26767C (77%), C26873T (2%). **MADRID:** C26645T (2,3%), T26767C (97,7%). **ANDALUCÍA:** T26767C (100%). **CATALUÑA:** T26767C (95,6%), A26786G (4,4%).

2021-Q3

CASTILLA Y LEÓN: T26543C (1,4%), C26549T (2,8%), C26681T (5,2%), T26767C (83,4%), A26786G (1,9%), G27014T (5,4%). **MADRID:** G26526T (1,4%), C26681T (2,7%), T26767C (92,3%), A26786G (1,1%), G27014T (2,5%). **CATALUÑA:** G26526T (5,4%), A26528T (1%), A26588G (1%), C26681T (0,7%), T26767C (81,2%), G26769A (1%), T26773C (1,4%), A26786G (2,3%), T26987C (1%), G27014T (0,8%), C27059T (2%), G27074A (0,6%), G27147T (1,5%).

orf6

2020

CASTILLA Y LEÓN: C27247T (27,3%), T27384C (81,8%). **MADRID:** T27205- (36,4%), C27213G (18,2%), A27307G (18,2%), T27384C (27,3%). **ANDALUCÍA:** C27297T (53,8%), G27364T (15,4%), T27384C (30,8%). **CATALUÑA:** T27215C (16,7%), C27247T (16,7%), C27297T (33,3%), A27302G (16,7%), T27384C (16,7%).

2021-Q1

CASTILLA Y LEÓN: C27323T (25%), T27384C (87,5%). **MADRID:** T27384C (100%). **ANDALUCÍA:** G27281T (25%), T27285C (75%). **CATALUÑA:** C27213T (14,3%), C27276T (14,3%), C27294T (14,3%), A27363C (14,3%), G27364T (14,3%), A27386- (28,6%).

2021-Q2

CASTILLA Y LEÓN: C27219A (17,4%), G27281T (8,7%), C27350T (47,8%), G27376T (13%), T27384C (17,4%). **MADRID:** C27297T (22,2%), G27376T (55,6%), T27384C (22,2%). **ANDALUCÍA:** C27263T (50%), T27285C (50%). **CATALUÑA:** T27342C (50%), A27344T (50%).

2021-Q3

CASTILLA Y LEÓN: A27259C (28,6%), T27384C (71,4%). **MADRID:** C27223G (80%), A27259C (20%). **CATALUÑA:** C27247T (38,5%), A27294G (11,5%), A27321T (26,9%), T27346C (11,5%), T27381C (11,5%).

orf7a

2020

CASTILLA Y LEÓN: C27494T (28,6%), C27600T (42,9%), G27746T (28,6%). **MADRID:** C27476T (25%), G27506T (8,3%), C27538T (8,3%), C27575T (8,3%), G27604T (8,3%), C27610T (8,3%), G27626T (8,3%), T27656C (8,3%), C27679T (8,3%), T27740C (8,3%). **ANDALUCÍA:** G27467T (16,7%), G27505T (16,7%), C27630T (16,7%), T27671C (16,7%), A27747G (16,7%), G27754T (16,7%). **CATALUÑA:** G27651T (100%).

2021-Q1

CASTILLA Y LEÓN: G27600T (33,3%), C27741T (40%), C27752T (26,7%). **MADRID:** G27463A (30%), T27501A (20%), C27527T (30%), C27532A (20%). **ANDALUCÍA:** G27506T (12,5%), G27750T (87,5%). **CATALUÑA:** C27513T (46,9%), G27517A (18,8%), C27577T (12,5%), G27750T (12,5%), C27752T (9,4%).

2021-Q2

CASTILLA Y LEÓN: C27513T (2,1%), C27577- (1,4%), A27579G (1,4%), T27638C (44,4%), C27739T (3,7%), G27750T (2,7%), C27752T (44,4%). **MADRID:** T27638C (47%), G27708A (1,5%), C27739T (3,8%), C27752T (47,8%). **ANDALUCÍA:** T27638C (36,4%), G27750T (27,3%), C27752T (36,4%). **CATALUÑA:** T27638C (50%), C27752T (50%).

2021-Q3

CASTILLA Y LEÓN: C27494T (1,2%), G27506T (1,4%), C27527T (1,2%), T27638C (47,6%), C27688T (0,7%), C27688- (0,1%), C27752T (47,8%).
MADRID: G27506T (0,8%), C27527G (0,2%), C27527T (1,1%), A27579- (0,8%), T27580- (0,8%), T27581- (0,8%), T27638C (46,1%), A27745G (0,8%), C27752T (48,7%).
CATALUÑA: G27420A (0,5%), A27446G (0,2%), T27465C (0,3%), T27484C (0,3%), C27494T (0,2%), G27506T (0,2%), A27507C (0,5%), T27511C (0,3%), C27527T (3,2%), T27529C (0,5%), G27541T (0,2%), G27561C (0,1%), G27561T (0,1%), C27575T (0,3%), C27577T (0,4%), C27590T (0,3%), T27638C (45%), C27649A (0,1%), C27649T (0,1%), C27688T (0,3%), G27731T (0,3%), A27743T (0,3%), A27745G (0,6%), C27752T (45,7%).

orf7b

2020

CASTILLA Y LEÓN: G27861T (100%).
MADRID: G27788T (11,1%), T27803C (11,1%), C27807T (11,1%), T27810C (11,1%), A27865T (22,2%), T27866A (11,1%), C27881T (11,1%), C27883T (11,1%).
ANDALUCÍA: G27777T (25%), C27804T (25%), T27817G (25%), C27881T (25%).
CATALUÑA: SIN MUTANTES.

2021-Q1

CASTILLA Y LEÓN: T27757C (50%), C27769T (50%).
MADRID: C27769T (50%), C27847T (50%).
ANDALUCÍA: C27800A (100%).
CATALUÑA: T27811C (50%), G27870T (50%).

2021-Q2

CASTILLA Y LEÓN: T27757C (2,5%), C27874T (97,5%).
MADRID: C27769T (1,5%), C27874T (98,5%).
ANDALUCÍA: C27874T (100%).
CATALUÑA: C27874T (100%).

2021-Q3

CASTILLA Y LEÓN: C27807T (2,1%), C27874T (97,9%).
MADRID: C27828T (1,5%), C27874T (98,5%).
CATALUÑA: C27874T (100%).

orf8

2020

CASTILLA Y LEÓN: C27944T (44,3%), C27972T (5,2%), G28048T (5,2%), A28111G (5,2%), T28144C (11,3%), G28193T (28,9%).
MADRID: C27944T (20,3%), T28144C (79,7%).
ANDALUCÍA: C27944T (35%), C27972T (2,8%), C27982T (15,9%), G28086T (7,9%), T28144C (38,3%).
CATALUÑA: C27944T (33,3%), G28077T (33,3%), T28144C (33,3%).

2021-Q1

CASTILLA Y LEÓN: C27944T (24,9%), C27972T (22,4%), C27982T (2,4%), G28048T (22,4%), C28093T (2,4%), A28095T (2,9%), A28111G (22,4%).
MADRID: C27944T (16,8%), C27972T (27,7%), G28048T (27,7%), A28111G (27,7%).
ANDALUCÍA: C27920T (1%), C27944T (3,8%), C27972T (29,6%), G28048T (29,5%), A28095T (6,3%), A28111G (29,3%), G28237T (0,5%).
CATALUÑA: C27972T (29,4%), C27982T (1%), G28001A (1%), G28048T (29,4%), C28093T (1,2%), A28095T (5,3%), A28111G (29,4%), C28112T (1,7%), G28167A (1%), C28253T (0,8%).

2021-Q2

CASTILLA Y LEÓN: G27948T (1,8%), C27972T (13,5%), C28005T (0,4%), C28013T (0,9%), G28048T (13,1%), G28090- (0,6%), T28091- (0,6%), T28092- (0,6%), C28093- (0,9%), T28094- (0,6%), A28095T (3,6%), A28111G (13,1%), G28237T (0,9%), G28248- (8,2%), A28249- (8,2%), T28250- (8,2%), T28251- (8,2%), T28252- (8,2%), C28253- (8,2%).
MADRID: C27925A (1,7%), C27972T (4,8%), C28005T (1,7%), G28048T (5,5%), G28086T (0,5%), C28093T (1,5%), A28095T (0,8%), A28111G (5%), G28209T (1,1%), G28248- (10,8%), A28249- (10,9%), T28250- (10,7%), T28251- (10,7%), T28252- (10,7%), C28253A (22,9%), A28254C (0,5%).
ANDALUCÍA: C27972T (20,8%), G28048T (20,8%), A28111G (20,8%), C28153T (0,8%), G28248- (6,2%), A28249- (6,2%), T28250- (6,2%), T28251- (6,2%), T28252- (6,2%), C28253- (6,2%).
CATALUÑA: G28248- (16,7%), A28249- (16,7%), T28250- (16,7%), T28251- (16,7%), T28252- (16,7%), C28253- (16,7%).

2021-Q3

CASTILLA Y LEÓN: C27975T (1%), T28007C (0,2%), C28045T (0,3%), G28086T (0,1%), C28093T (0,3%), G28248- (16,3%), A28249- (16,3%), T28250- (16,3%), T28251- (16,3%), T28252- (16,3%), C28253- (16,3%).
MADRID: C27925A (0,3%), C28005T (0,3%), G28086T (0,3%), G28248- (16,4%), A28249- (16,4%), T28250- (16,4%), T28251- (16,4%), T28252- (16,4%), C28253A (0,7%), C28253- (16,4%).
CATALUÑA: G27916A (0,1%), G27916T (0,1%), C27944T (0,1%), C27945T (0,1%), G28001T (0,2%), T28007C (0,4%), G28027A (0,01%), G28027C (0,01%), G28027T (0,1%), G28086T (1%), C28093T (0,3%), G28116A (0,3%), T28156C (0,1%), G28209T (0,2%), T28245- (0,01%), T28245G (0,9%), G28248- (16%), A28249- (16%), T28250- (16%), T28251- (16%), T28252- (16%), C28253- (16%).

n

2020

CASTILLA Y LEÓN: C28657T (6,9%), C28863T (4,7%), C28932T (76,3%), G29235T (12,1%).
MADRID: C28333T (4,2%), C28657T (17,3%), C28863T (17,3%), G28881A (12,5%), G28882A (11,3%), G28883C (11,3%), C28932T (22%), G29402T (3,6%).
ANDALUCÍA: C28657T (9,7%), C28833T (1,3%), C28863T (9,9%), G28881A (4,1%), G28882A (3,9%), G28883C (3,9%), C28932T (34,2%), G28975C (2,7%), C29366T (1,7%), C29386T (3,2%), G29398A (2,1%), G29399A (2,7%), G29402T (17,6%), G29425T (3%).
CATALUÑA: G28881A (32,7%), G28882A (32,7%), G28883C (34,5%).

2021-Q1

CASTILLA Y LEÓN: G28280C (10,8%), A28281T (10,8%), T28282A (10,8%), C28849T (1,4%), G28881A (10,8%), G28882A (11,1%), G28883C (10,8%), C28932T (15,5%), C28977T (10,8%), C29451T (7,1%).
MADRID: G28280C (11,4%), A28281T (11,4%), T28282A (11,4%), G28690T (1,4%), G28881A (11,8%), G28882A (11,8%), G28883C (11,8%), C28887T (2,6%), C28932T (10%), C28977T (11,2%), G29402T (3,1%), G29513T (2%).
ANDALUCÍA: G28280C (12%), A28281T (12%), T28282A (12%), G28346T (0,5%), C28651T (0,6%), C28869T (0,4%), G28881A (12%), G28882A (12%), G28883C (12%), C28932T (6,4%), G28975C (1%), C28977T (12%), G29398A (1,6%), G29399A (1%), G29402T (4,3%).
CATALUÑA: G28280C (12,9%), A28281T (12,9%), T28282A (12,9%), C28512G (0,4%), A28877T (0,7%), G28878C (0,7%), G28881A (13,6%), G28882A (13,6%), G28883C (13,6%), C28887T (0,9%), C28932T (3,3%), A28951G (0,7%), C28977T (12,8%), C29366T (0,7%), G29513T (0,4%).

2021-Q2

CASTILLA Y LEÓN: G28280C (8,7%), A28281T (8,7%), T28282A (10,2%), A28299T (1,8%), A28461G (6,5%), A28733G (0,7%), C28849T (1,7%), G28881A (16,7%), G28882A (10,2%), G28883C (10,2%), G28884C (0,2%), G28916T (5,3%), C28977T (10,3%), G29140T (0,3%), C29149T (0,3%), C29250T (0,2%), C29253T (0,3%), C29296T (0,3%), G29402T (6,5%), G29422A (0,2%), C29445T (0,6%).
MADRID: G28280C (3,7%), A28281T (3,7%), T28282A (3,7%), C28291T (0,4%), A28299T (5,5%), A28461G (9,9%), C28657T (0,4%), A28877T (0,4%), G28878C (0,4%), G28881T (22,5%), G28882A (4%), G28883C (4%), C28887T (1,6%), G28899T (0,4%), G28903T (0,4%), G28916T (14,4%), C28977T (3,5%), C29253T (1,7%), C29311T (0,6%), G29402T (18,3%), A29406T (0,4%).
ANDALUCÍA: G28280C (12,1%), A28281T (12,1%), T28282A (12,1%), G29402T (3,6%).
CATALUÑA: A28299T (13,4%), A28461G (21,3%), G28881T (15,7%), G28882A (12,1%), G28883C (12,1%), G28916T (2,7%), C28977T (12,1%), G29402T (21,3%).

2021-Q3

CASTILLA Y LEÓN: A28299T (8,3%), A28309G (0,2%), C28311T (0,4%), A28363- (0,3%), G28364- (0,3%), A28365- (0,3%), A28366- (0,3%), C28367- (0,3%), G28368- (0,3%), C28369- (0,3%), A28370- (0,3%), A28461G (20,9%), C28486A (0,4%), C28507T (0,6%), C28531T (0,4%), G28703C (0,3%), G28857C (0,1%), G28857T (0,4%), C28866T (0,4%), G28881A (0,4%), G28881T (21%), G28882A (0,4%), G28883C (0,4%), G28916T (19,2%), C28957T (0,2%), G29179T (1%), G29195T (0,2%), C29253T (0,2%), G29260T (0,2%), C29296T (0,4%), T29335C (0,4%), G29392T (0,2%), G29402T (20,9%).
MADRID: A28299T (10,6%), A28461G (21,7%), G28881A (0,3%), G28881T (22,4%), C28887A (0,3%), C28887T (0,3%), G28916T (21,6%), G29402T (22,7%).
CATALUÑA: A28299T (6,9%), G28378A (0,1%), G28378T (2,3%), G28396A (0,8%), A28461G (29,5%), C28610A (0,4%), G28881T (29,5%), G28916T (29%), G28922T (0,6%), C29077T (0,6%), C29119G (0,3%).

orf10

2020

CASTILLA Y LEÓN: G29645T (100%).
MADRID: G29645T (100%).
ANDALUCÍA: C29614T (3,4%), T29619A (1,9%), A29630T (1,9%), G29645T (92,7%).
CATALUÑA: G29645T (100%).

2021-Q1

CASTILLA Y LEÓN: G29645T (100%).
MADRID: C29614T (17,9%), G29645T (82,1%).
ANDALUCÍA: G29645T (100%).
CATALUÑA: C29614T (12,2%), G29645T (87,8%).

2021-Q2

CASTILLA Y LEÓN: C29614T (33,3%), C29640T (33,3%), C29642T (33,3%).
MADRID: C29563T (9,1%), C29580T (9,1%), C29585G (18,2%), T29589C (18,2%), C29592T (9,1%), C29614T (9,1%), C29640T (18,2%), G29645T (9,1%).
ANDALUCÍA: T29589C (50%), C29614T (50%).
CATALUÑA: T29589C (33,3%), G29639C (33,3%), G29645T (33,3%).

2021-Q3

CASTILLA Y LEÓN: G29562A (5%), C29585T (5%), C29614T (20%), G29616C (10%), C29627T (5%), G29645T (10%), G29648T (25%), C29666T (20%).
MADRID: T29564C (11,1%), T29566C (11,1%), C29592T (44,4%), C29642T (11,1%), A29647G (11,1%), G29648T (11,1%).
CATALUÑA: (C29572T (2,6%), C29572- (2,6%), C29614T (5,3%), G29628T (2,6%), G29628A (10,5%), C29642T (7,9%), G29648T (52,6%), A29649G (5,3%), C29659T (5,3%), G29674T (5,3%).

ANEXO II. Mutaciones de SARS-CoV-2 relevantes destacadas entre las comunidades de Castilla y León, Madrid, Andalucía y Cataluña. La tabla indica respectivamente, desde el inicio de la primera hasta la última columna: el nombre del gen afectado, el nombre de la mutación, el cambio aminoacídico que produce, la fecha en la que se ha encontrado la mutación, nombre de la región afectada del gen o proteína, tipo de mutación (DEL = delección; TS = transición; TV = transversión; INS = inserción; CC = cambio de codón completo) y su subtipo (DEL = delección; SY = cambio sinónimo; NS = cambio no sinónimo; STOP = codón de parada). SNP en subrayado; * = STOP. RdRp = NSP12

Gen	Mutación	Cambio aminoacídico	Fecha detectada	Región	Tipo	Subtipo
<i>orf1ab</i>	686-694del	KSF141-143Δ	2021-Q3	NSP1	DEL	DEL
	<u>G1820A</u>	G519S	2020, 2021-Q1, Q3	NSP2	TS	NS
	<u>C3037T</u>	F924F	2020, 2021-Q1, Q2, Q3	NSP3	TS	SY
	<u>C4543T</u>	T1426T	2020, 2021-Q1, Q2, Q3	NSP3	TS	SY
	<u>C5388A</u>	A1708D	2020, 2021-Q1, Q2	NSP3	TV	NS
	<u>C5986T</u>	F1907F	2021-Q1, Q2	NSP4	TS	SY
	<u>C6402T</u>	P2046L	2021-Q2, Q3	NSP4	TS	NS
	<u>C7124T</u>	P2287S	2021-Q2, Q3	NSP4	TS	NS
	<u>G9053T</u>	V2930L	2021-Q2, Q3	NSP4	TV	NS
	<u>C9891T</u>	A3209V	2020, 2021-Q2, Q3	NSP4	TS	NS
	<u>C10029T</u>	T3255I	2021-Q1, Q2, Q3	NSP4	TS	NS
	<u>A10323G</u>	K3353R	2021-Q1, Q2, Q3	NSP5	TS	NS
	<u>C10449A/C10449T</u>	P3395L/P3395H	2021-Q3	NSP5	TS/TV	NS
	<u>G11083T</u>	L3606F	2020, 2021-Q1, Q3	NSP6	TV	NS
	<u>A11201G</u>	T3646A	2021-Q2, Q3	NSP6	TS	NS
	11288-11296del	SGF3675-3677Δ	2021-Q1, Q2, Q3	NSP6	DEL	DEL
	<u>A11332G</u>	V3689V	2021-Q2, Q3	NSP6	TS	SY
	<u>C11514T</u>	T3750I	2021-Q1, Q2, Q3	NSP6	TS	NS
	<u>C12970T</u>	N4235N	2020, 2021-Q1, Q2	NSP9	TS	SY
	<u>C14408T</u>	P4715L	2020, 2021-Q1, Q2, Q3	NSP12	TS	NS
	<u>G15451A</u>	G5063S	2021-Q2, Q3	NSP12	TS	NS
	<u>C15952A</u>	L5230I	2021-Q2, Q3	NSP12	TV	NS
	<u>C16466T</u>	P5401L	2021-Q2, Q3	NSP13	TS	NS
	<u>C18877T</u>	L6206L	2020, 2021-Q1, Q2, Q3	NSP14	TS	SY
	<u>C19220T</u>	A6319V	2021-Q2, Q3	NSP14	TS	NS
	<u>A20268G</u>	L6668L	2020, 2021-Q1	NSP15	TS	SY
	<u>G21255C</u>	A6997A	2020, 2021-Q1	NSP16	TV	SY
<i>s (spike)</i>	<u>C21618G</u>	T19R	2021-Q2, Q3	S1(NTD)	TV	NS
	21765-21770del	HV69-70Δ	2020, 2021-Q1, Q2	S1(NTD)	DEL	DEL
	<u>C21846T</u>	T95I	2021-Q1, Q2, Q3	S1(NTD)	TS	NS
	<u>G21987A</u>	G142D	2021-Q2, Q3	S1(NTD)	TS	NS
	21991-21993del	Y144Δ	2020, 2021-Q1, Q2, Q3	S1(NTD)	DEL	DEL
	22029-22034del	EF156-157Δ	2021-Q2, Q3	S1(NTD)	DEL	DEL
	<u>C22227T</u>	A222V	2020, 2021-Q1, Q2, Q3	S1(NTD)	TS	NS
	<u>T22917G</u>	L452R	2021-Q2, Q3	S1(RBD)	TV	NS
	<u>C22995A</u>	T478K	2021-Q2, Q3	S1(RBD)	TV	NS
	<u>G23012A/G23012C</u>	E484K/E484Q	2021-Q1, Q2, Q2 / Q3	S1(RBD)	TS/TV	NS
	<u>A23063T</u>	N501Y	2020, 2021-Q1, Q2, Q3	S1(RBD)	TV	NS
	<u>C23271A</u>	A570D	2020, 2021-Q1, Q2	S1	TV	NS
	<u>A23403G</u>	D614G	2020, 2021-Q1, Q2, Q3	S1	TS	NS
	<u>C23525T</u>	H655Y	2021-Q1, Q3	S1/S2	TS	NS
	<u>C23604A/C23604G</u>	P681H/P681R	2021-Q1, Q2 / Q2, Q3	S1/S2	TV	NS
<i>orf3a</i>	<u>C23709T</u>	T716I	2020, 2021-Q1, Q2	S2	TS	NS
	<u>T24506G</u>	S982A	2020, 2021-Q1, Q2	S2	TV	NS
	<u>G24914C</u>	D1118H	2020, 2021-Q1, Q2	S2	TV	NS
	<u>C25469T</u>	S26L	2021-Q2, Q3	N/A	TS	NS
	<u>G25563T</u>	Q57H	2020, 2021-Q1, Q2	N/A	TV	NS
	<u>C25626T</u>	H78Y	2021-Q1, Q2, Q3	N/A	TS	NS
	<u>ins25702-25704CCC</u>	ins104P	2021-Q3	N/A	INS	INS
<i>orf3b</i>	<u>G25855T</u>	D155Y	2021-Q1, Q3	N/A	TV	NS
	<u>G25979T</u>	G196V	2020	N/A	TV	NS
	<u>C25844T</u>	T27I	2021-Q1, Q2, Q3	N/A	TS	NS
	<u>G25855T</u>	D31Y	2021-Q1, Q2, Q3	N/A	TV	NS
	<u>C25872T</u>	Y36Y	2021-Q2, Q3	N/A	TS	SY
	<u>G25906C</u>	G48R	2020, 2021-Q1, Q3	N/A	TV	NS
	<u>G25979T</u>	G72V	2020	N/A	TS	NS
	<u>G25996T</u>	V78L	2020, 2021-Q1, Q3	N/A	TS	NS
	<u>C26060T</u>	T99I	2021-Q2, Q3	N/A	TS	NS

<i>e</i>	G26268T	E8D	2020	N/A	TV	NS
	C26270T	T9I	2021-Q1, Q2, Q3	N/A	TS	NS
	C26408T	S55F	2021-Q2, Q3	N/A	TS	NS
	T26409C	S55S	2021-Q3	N/A	TS	SY
	G26428T	S62F	2021-Q2, Q3	N/A	TV	NS
	T26442A	N66K	2021-Q3	N/A	TV	SY
	C26444T	S68F	2021-Q3	N/A	TS	NS
	C26456T	P71L	2021-Q1, Q2	N/A	TS	NS
	C26464T	L74L	2021-Q1, Q2, Q3	N/A	TS	SY
	G26526T	A2S	2021-Q1, Q2, Q3	N/A	TV	NS
<i>m</i>	C26681T	F53F	2020, 2021-Q3	N/A	TS	SY
	T26767C	I82T	2021-Q1, Q2, Q3	N/A	TS	NS
	A26786G	G89G	2021-Q2, Q3	N/A	TS	SY
	C26801G	L93L	2020, 2021-Q1	N/A	TV	SY
	G27014T	P165P	2021-Q3	N/A	TV	SY
	C27223G	Q8E	2021-Q3	N/A	TV	NS
<i>orf6</i>	C27247T	L16L	2020, 2021-Q3	N/A	TS	SY
	A27259C	R20R	2021-Q3	N/A	TV	SY
	T27285C	N28N	2021-Q1, Q2	N/A	TS	SY
	C27297T	I32I	2020, 2021-Q2	N/A	TS	SY
	A27321T	L40F	2021-Q3	N/A	TV	NS
	G27364T	E55*	2020, 2021-Q1	N/A	TV	STOP
	G27376T	E59*	2021-Q2	N/A	TV	STOP
	T27384C	D61L	2020, 2021-Q1, Q2, Q3	N/A	TS	SY
	G27506T	G38V	2020, 2021-Q1, Q3	N/A	TV	NS
<i>orf7a</i>	T27638C	V82A	2021-Q1, Q2, Q3	N/A	TS	NS
	C27688T	P99S	2021-Q3	N/A	TS	NS
	C27741T	L116L	2021-Q1	N/A	TS	SY
	G27750T	K119N	2021-Q1, Q2	N/A	TV	NS
	C27752T	T120I	2021-Q1, Q2, Q3	N/A	TS	NS
<i>orf7b</i>	C27769T	S5L	2021-Q1, Q2	N/A	TS	NS
	C27807T	L18L	2020, 2021-Q3	N/A	TS	SY
	C27874T	T40I	2021-Q2, Q3	N/A	TS	NS
<i>orf8</i>	C27944T	H17H	2020, 2021-Q1, Q3	N/A	TS	SY
	C27972T	Q27*	2020, 2021-Q1, Q2	N/A	TS	STOP
	G28048T	R52I	2020, 2021-Q1, Q2	N/A	TV	NS
	G28086T	A65S	2020, 2021-Q2, Q3	N/A	TV	NS
	C28093T	S67F	2021-Q1, Q2, Q3	N/A	TS	NS
	A28095T	K68*	2021-Q1, Q2	N/A	TV	STOP
	A28111G	Y73C	2020, 2021-Q1, Q2	N/A	TS	NS
	T28144C	L84S	2020	N/A	TS	NS
	28248-28253del	119-120Δ	2021-Q2, Q3	N/A	DEL	DEL
<i>n</i>	GAT28280-28282CTA	D3L	2021-Q1, Q2	<i>orf9a</i>	CC	NS
	A28299T	Q9L/ORF9B (S6C)	2021-Q2, Q3	<i>orf9b</i>	TV	NS
	A28461G	D63G/ORF9B (T60A)	2021-Q2, Q3	<i>orf9b</i>	TS	NS
	C28657T	D128D	2020, 2020-Q2	<i>orf9a</i>	TS	SY
	C28863T	S197L/ORF9C (Q44*)	2020	<i>orf9c</i>	TS	NS
	G28881T	R203M/ORF9C (G50W)	2021-Q2, Q3	<i>orf9c</i>	TV	NS
	G28881A	R203K/ORF9C (G50K)	2020, 2021-Q1, Q2, Q3	<i>orf9c</i>	TS	NS
	G28882A		2021-Q2, Q3	<i>orf9c</i>	TV	NS
	G28881T		2021-Q2, Q3	<i>orf9c</i>	TS	
	G28882A		2021-Q2, Q3		TV	
	G28883C	G204R/ORF9C (G50G)	2020, 2021-Q1, Q2, Q3	<i>orf9c</i>	TV	NS/SY
	C28887T	T205I/ORF9C (L52F)	2021-Q1, Q2, Q3	<i>orf9c</i>	TS	NS
	G28916T	G215C/ORF9C (A61A)	2021-Q2, Q3	<i>orf9c</i>	TV	NS/SY
	C28932T	A220V/ORF9C (L67F)	2020, 2021-Q1	<i>orf9c</i>	TS	NS/SY
	C28977T	S235F	2021-Q1, Q2	<i>orf9a</i>	TS	NS
	C29253T	S327L	2021-Q2, Q3	<i>orf9a</i>	TS	NS
<i>orf10</i>	G29402T	D377Y	2020, 2021-Q1, Q2, Q3	N/A	TV	NS
	T29589C	F11S	2021-Q2	N/A	TS	NS
	C29614T	C19C	2020, 2021-Q1, Q2, Q3	N/A	TS	SY
	G29645T	V30L	2020, 2021-Q1, Q2, Q3	N/A	TV	NS
	G29648T	D31Y	2021-Q3	N/A	TV	NS

ANEXO III – Mutaciones relevantes y predominantes de SARS-CoV-2 presentes desde el inicio hasta el final de cada intervalo temporal que aparecen entre las comunidades de Castilla y León, Madrid, Andalucía y Cataluña en 2020, los 3 cuatrimestres de 2021, y en ambos años. La primera columna indica el gen afectado, la segunda el nombre de la mutación y la tercera indica el porcentaje que ocupa la mutación en la respectiva posición respecto a la cantidad de mutantes de ese intervalo temporal.

MUTACIONES RELEVANTES Y COMUNES DE 2020		
<i>orf1ab</i>	C3037T	93%
	C14408T	92,08%
<i>s (spike)</i>	A23403G	99%
MUTACIONES RELEVANTES Y COMUNES DE 2021-Q1		
<i>orf1ab</i>	C3037T	100%
	C14408T	99,9%
<i>s (spike)</i>	A23403G	100%
<i>m</i>	C26801G	85,3%
MUTACIONES RELEVANTES Y COMUNES DE 2021-Q2		
<i>orf1ab</i>	C3037T	99,6%
	C14408T	99,1%
<i>s (spike)</i>	A23403G	100%
MUTACIONES RELEVANTES Y COMUNES DE 2021-Q3		
<i>orf1ab</i>	C3037T	97,8%
	G4181T	93,1%
	C6402T	92,2%
	C7124T	93,1%
	G9053T	92,4%
	C10029T	94,2%
	A11201G	92,9%
	A11332G	93,1%
	C14408T	96,9%
	G15451A	98%
	C16466T	97,5%
	C19220T	92,9%
<i>s (spike)</i>	C21618G	92,4%
	G21987A	95,9%
	C22995A	98,3%
	A23403G	100%
	C23604G	97,7%
<i>orf3a</i>	G24410A	97,8%
	C25469T	99%
<i>m</i>	T26767C	100%
<i>orf7a</i>	T27638C	97,3%
	C27752T	99,7%
<i>orf7b</i>	C27874T	99,4%
<i>orf8</i>	28248-28253del	98,3%
<i>n</i>	A28461G	96,4%
	G28881T	97,6%
	G28916T	93%
MUTACIONES RELEVANTES Y COMUNES EN 2020 Y 2021		
<i>orf1ab</i>	C3037T	98%
	C14408T	97%
<i>s (spike)</i>	A23403G	100%