

Los genomas del SARS-CoV-2

Mireia Coscollá Devís

Institute for Integrative Systems Biology I²SysBio (Universitat de València-CSIC)



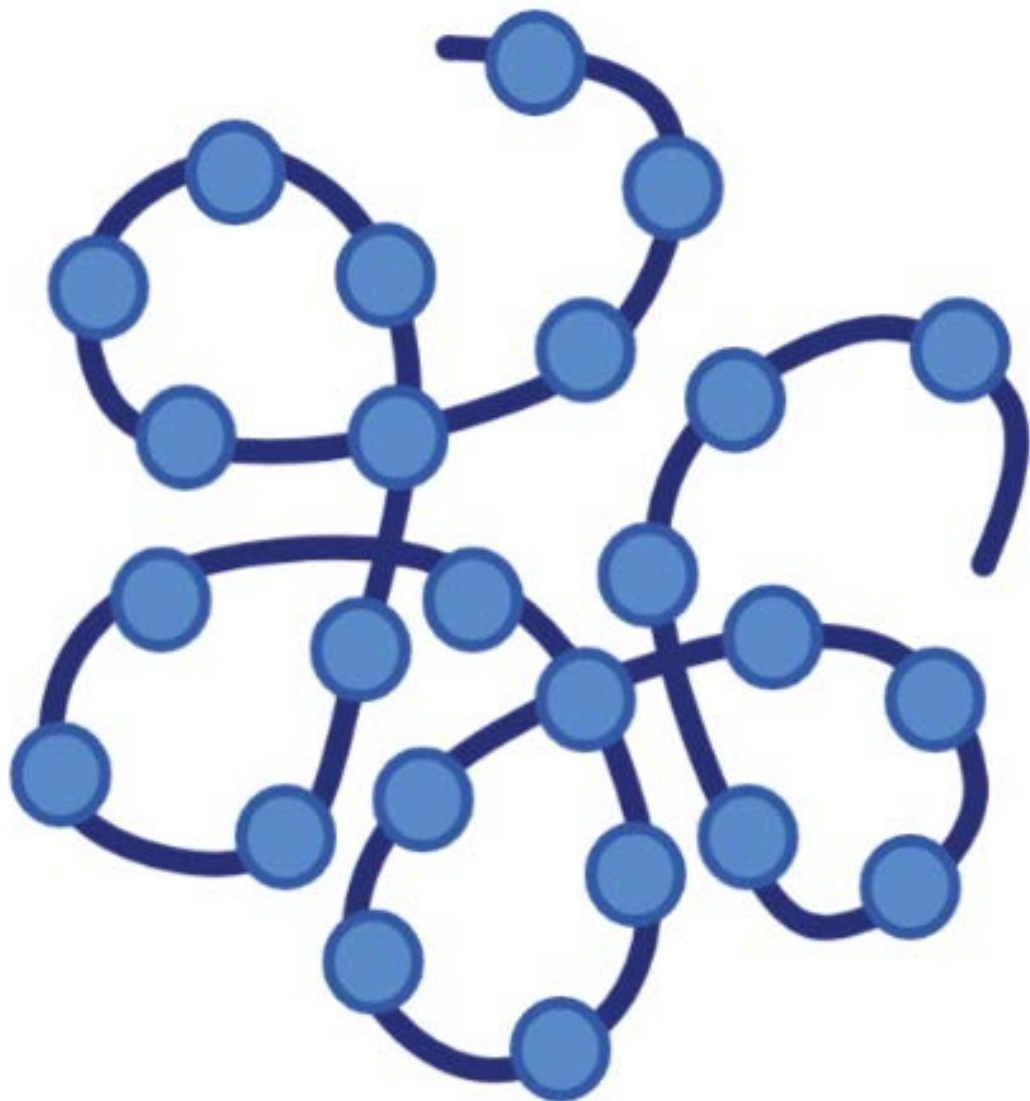


Institute for
Integrative Systems
Biology I²SysBio

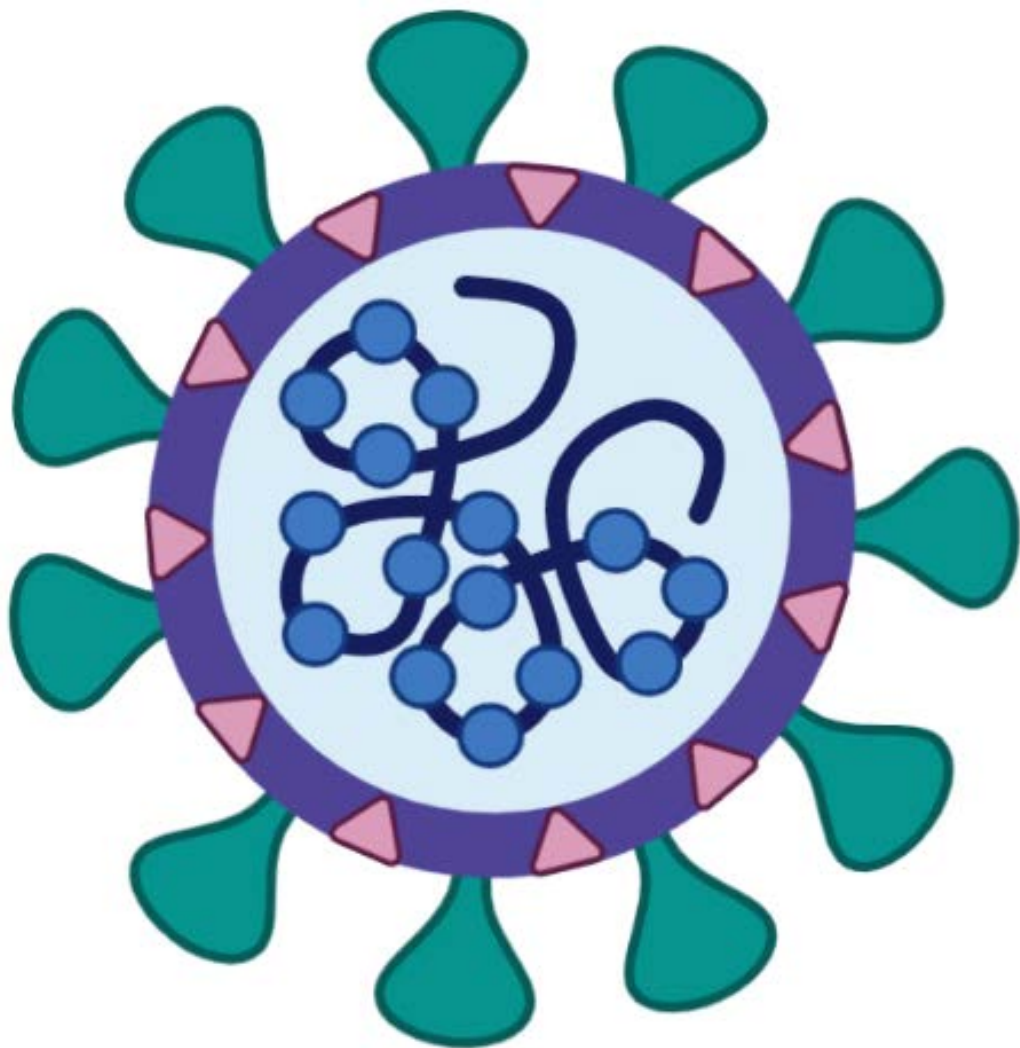
València, Spain



SARS-CoV-2: ARN envuelto en proteínas



SARS-CoV-2: ARN envuelto en proteínas



SARS- CoV 2 Structure. Contributed by Rohan Bir Singh, MD;
Made with Biorender.com

A pneumonia outbreak associated with a

<https://doi.org/10.1038/s41586-020-2012-7>

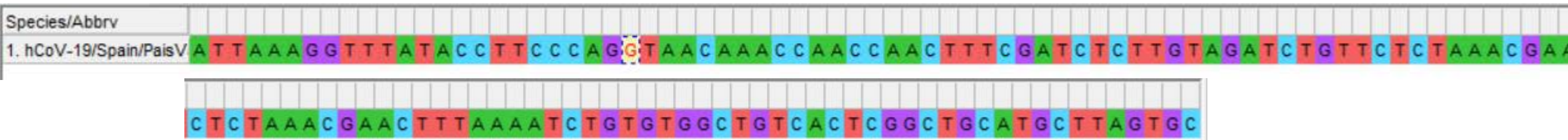
Received: 20 January 2020

Accepted: 29 January 2020

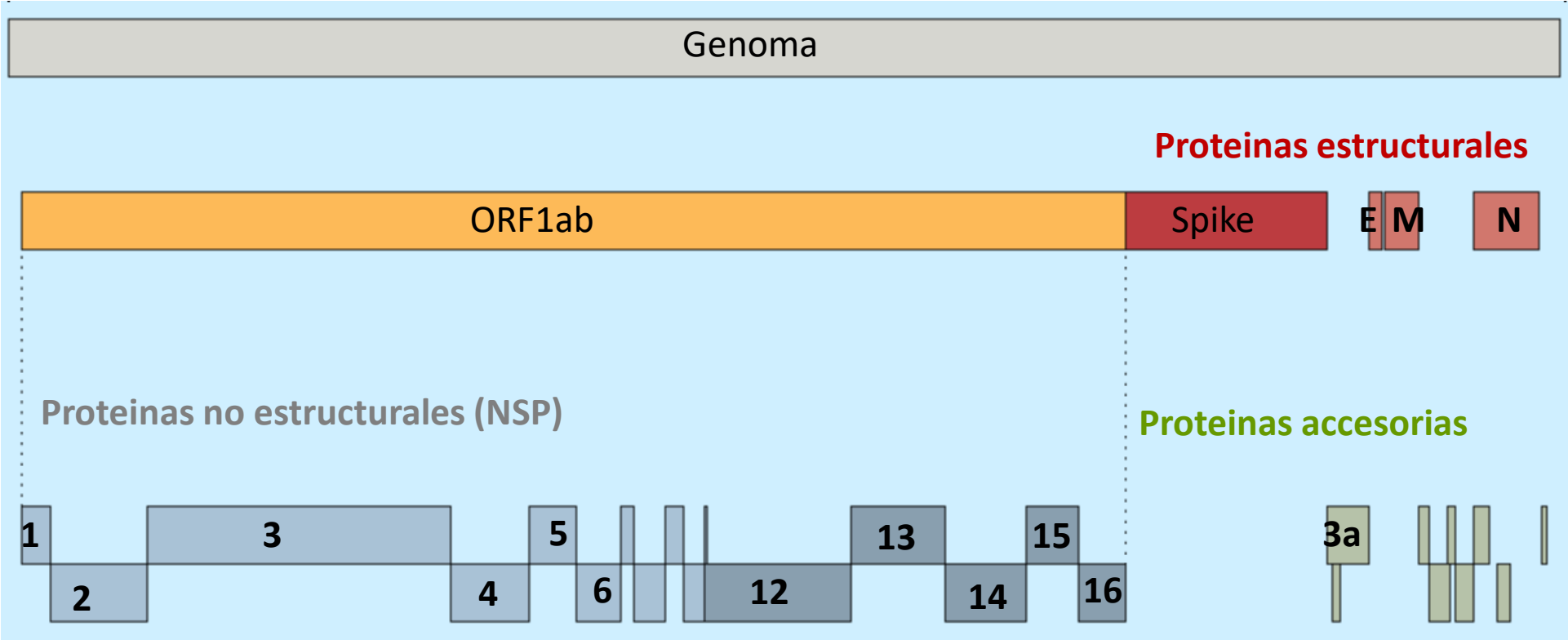
Published online: 3 February 2020

Peng Zhou^{1,5}, Xing-Lou Yang^{1,5}, Xian-Guang Wang^{2,5}, Ben Hu¹, Lei Zhang¹, Wei Zhang¹, Hao-Rui Si^{1,3}, Yan Zhu¹, Bei Li¹, Chao-Lin Huang², Hui-Dong Chen², Jing Chen^{1,3}, Yun Luo^{1,3}, Hua Guo^{1,3}, Ren-Di Jiang^{1,3}, Mei-Qin Liu^{1,3}, Ying Chen^{1,3}, Xu-Rui Shen^{1,3}, Xi Wang^{1,3}, Xiao-Shuang Zheng^{1,3}, Kai Zhao^{1,3}, Quan-Jiao Chen¹, Fei Deng¹, Lin-Lin Liu⁴, Bing Yan¹, Fa-Xian Zhan⁴, Yan-Yi Wang¹, Geng-Fu Xiao¹ & Zheng-Li Shi¹✉

La secuenciación de ARN metagenómica de una muestra de líquido de lavado broncoalveolar del paciente identificó una nueva cepa de virus de ARN de la familia Coronaviridae, que se denomina aquí coronavirus 'WH-Human 1' (y también se conoce como '2019-nCoV'). El análisis filogenético del genoma viral completo (29,903 nucleótidos) reveló que el virus está estrechamente relacionado (similitud de nucleótidos del 89.1%) con un grupo de coronavirus similares al SARS (género Betacoronavirus, subgénero, sarbecovirus) que se habían encontrado previamente en murciélagos en China.



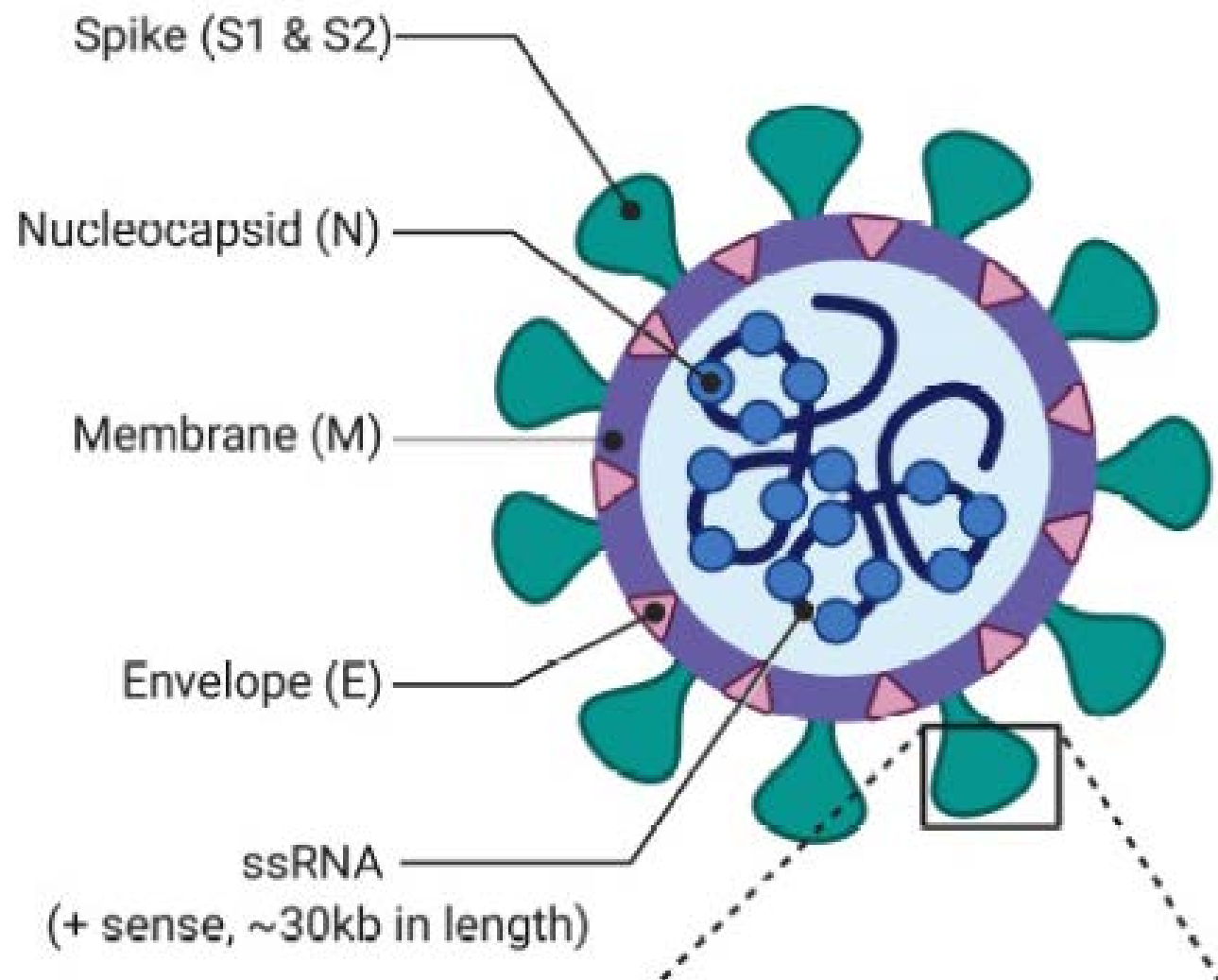
Organización genómica de SARS-CoV-2



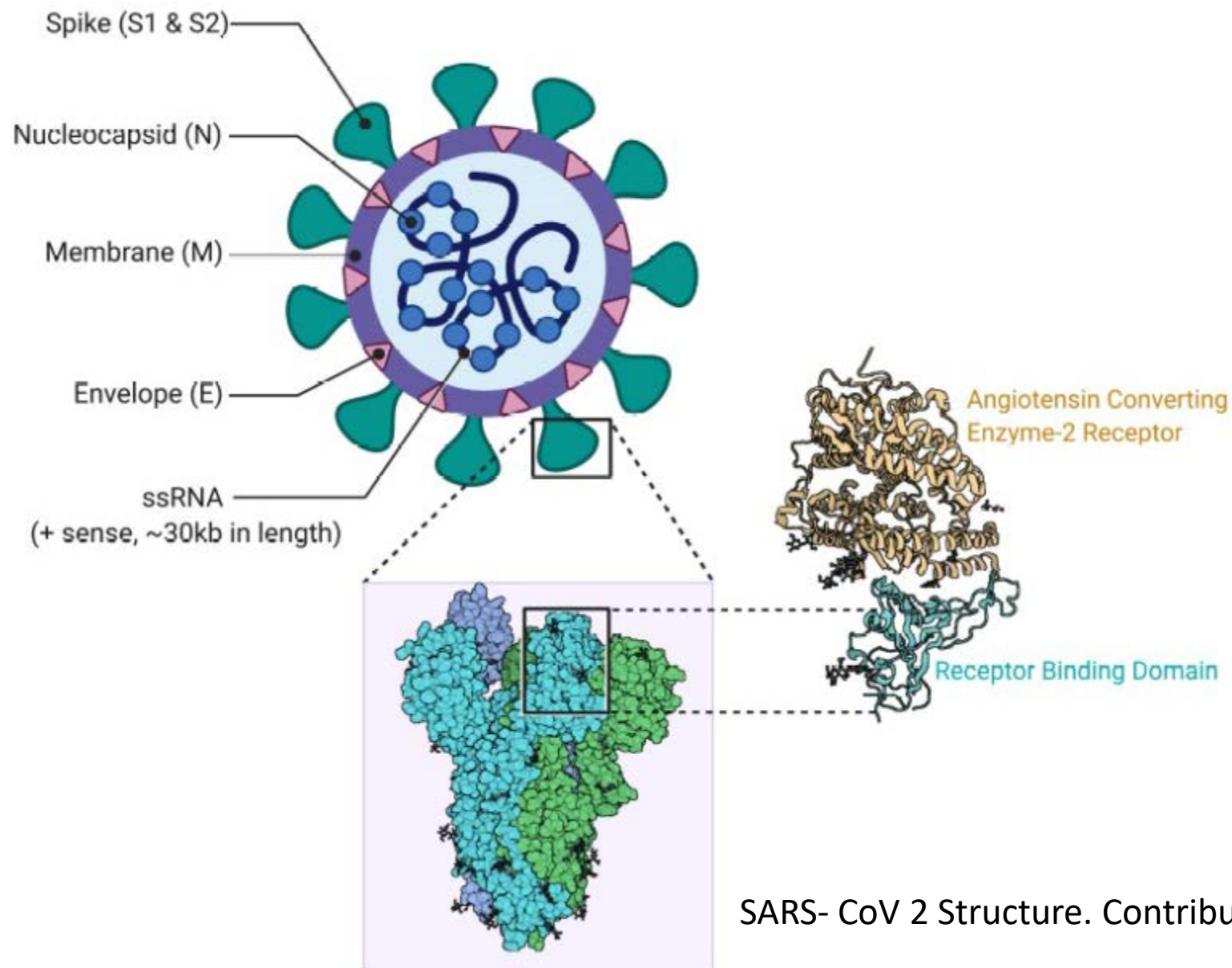
NSP1 sabotea la célula
NSP2 ¿?
NSP3 Proteasa
NSP4 DSP6 burbujas
NSP7 NSP8
asisten polimerasa

NSP12 RNA pol RNA dep
NSP9 abre canales nucleo
NSP10 NSP16 camuflaje genético
NSP13 Helicasa
NSP14 Corrección errores
NSP15 Limpia

Estructura SARS-CoV2



Estructura SARS-CoV2



SARS- CoV 2 Structure. Contributed by Rohan Bir Singh, MD;
Made with Biorender.com

Los primeros genomas

A pneumonia outbreak associated with a new coronavirus of probable bat origin

<https://doi.org/10.1038/s41586-020-2012-7>

Received: 20 January 2020

Accepted: 29 January 2020

Published online: 3 February 2020

Peng Zhou^{1,5}, Xing-Lou Yang^{1,5}, Xian-Guang Wang^{2,5}, Ben Hu¹, Lei Zhang¹, Wei Zhang¹, Hao-Rui Si^{1,3}, Yan Zhu¹, Bei Li¹, Chao-Lin Huang², Hui-Dong Chen², Jing Chen^{1,3}, Yun Luo^{1,3}, Hua Guo^{1,3}, Ren-Di Jiang^{1,3}, Mei-Qin Liu^{1,3}, Ying Chen^{1,3}, Xu-Rui Shen^{1,3}, Xi Wang^{1,3}, Xiao-Shuang Zheng^{1,3}, Kai Zhao^{1,3}, Quan-Jiao Chen¹, Fei Deng¹, Lin-Lin Liu⁴, Bing Yan¹, Fa-Xian Zhan⁴, Yan-Yi Wang¹, Geng-Fu Xiao¹ & Zheng-Li Shi¹✉

A new coronavirus associated with human respiratory disease in China

<https://doi.org/10.1038/s41586-020-2008-3>

Received: 7 January 2020

Accepted: 28 January 2020

Fan Wu^{1,7}, Su Zhao^{2,7}, Bin Yu^{3,7}, Yan-Mei Chen^{1,7}, Wen Wang^{4,7}, Zhi-Gang Song^{1,7}, Yi Hu^{2,7}, Zhao-Wu Tao², Jun-Hua Tian³, Yuan-Yuan Pei¹, Ming-Li Yuan², Yu-Ling Zhang¹, Fa-Hui Dai¹, Yi Liu¹, Qi-Min Wang¹, Jiao-Jiao Zheng¹, Lin Xu¹, Edward C. Holmes^{1,5} & Yong-Zhen Zhang^{1,4,6}✉

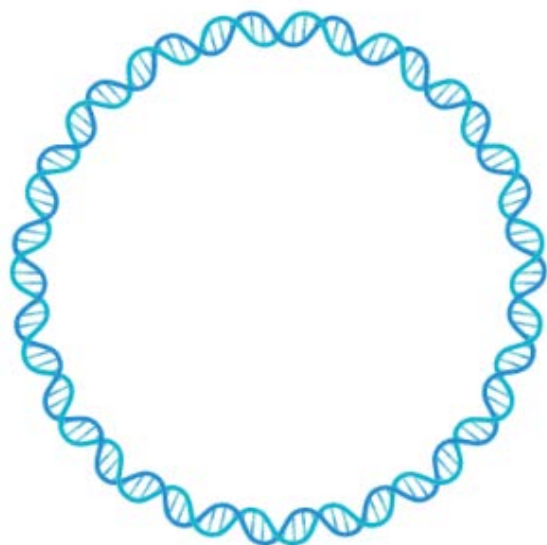
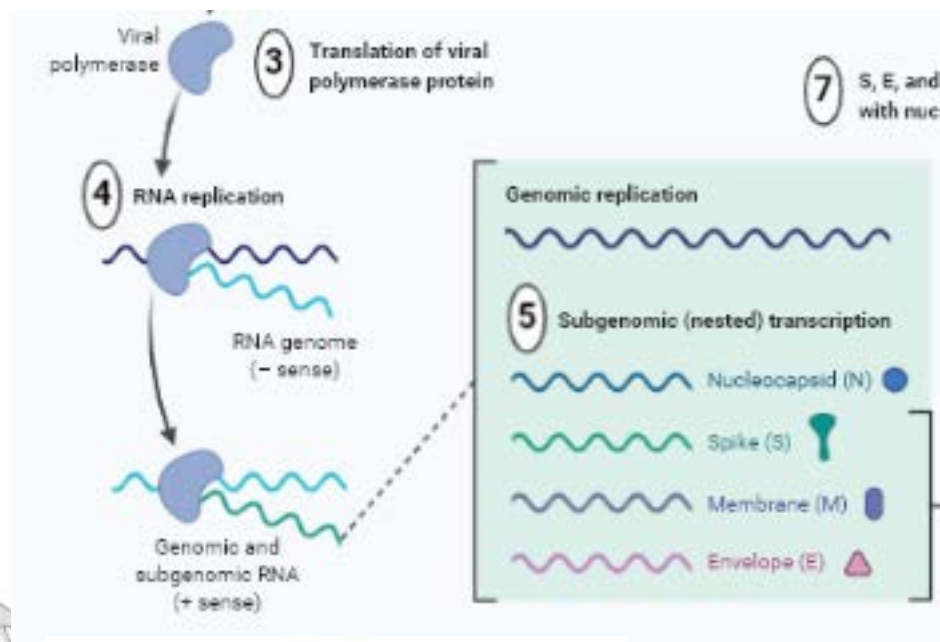
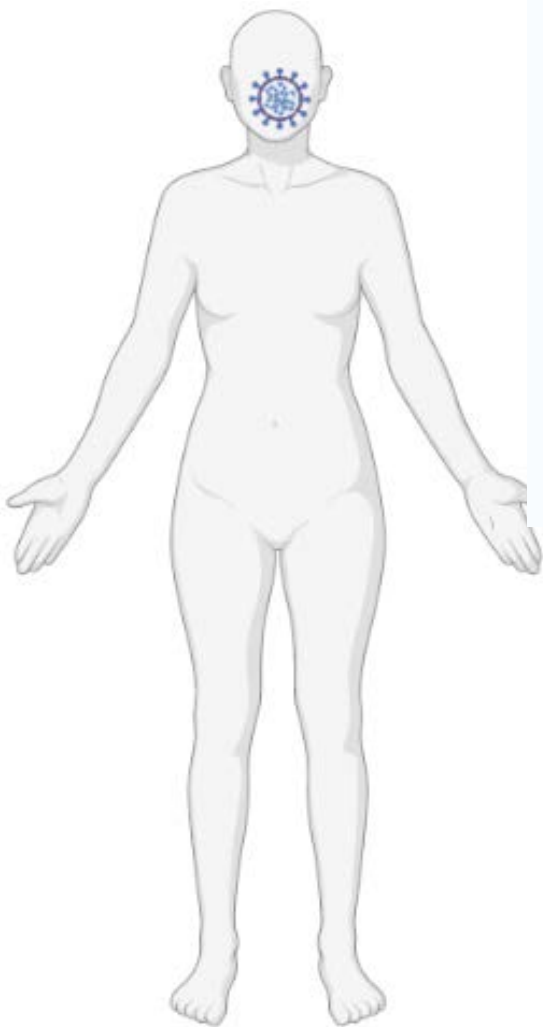
Se obtuvieron secuencias del genoma completas de cinco pacientes en una etapa temprana del brote. Las secuencias son casi idénticas y comparten una identidad de secuencia del 79,6% con el SARS-CoV.

Genomas SARS disponibles

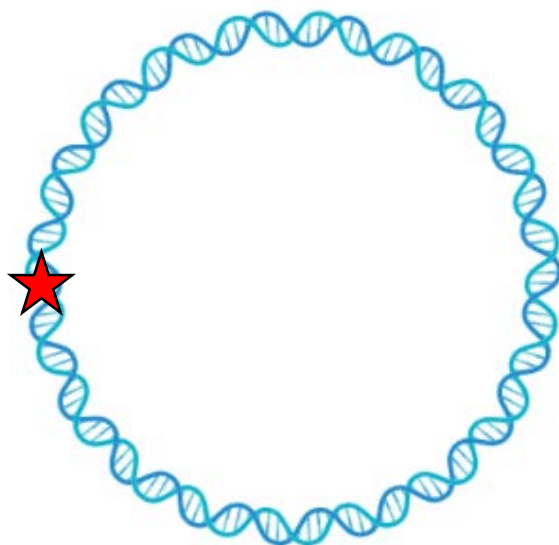
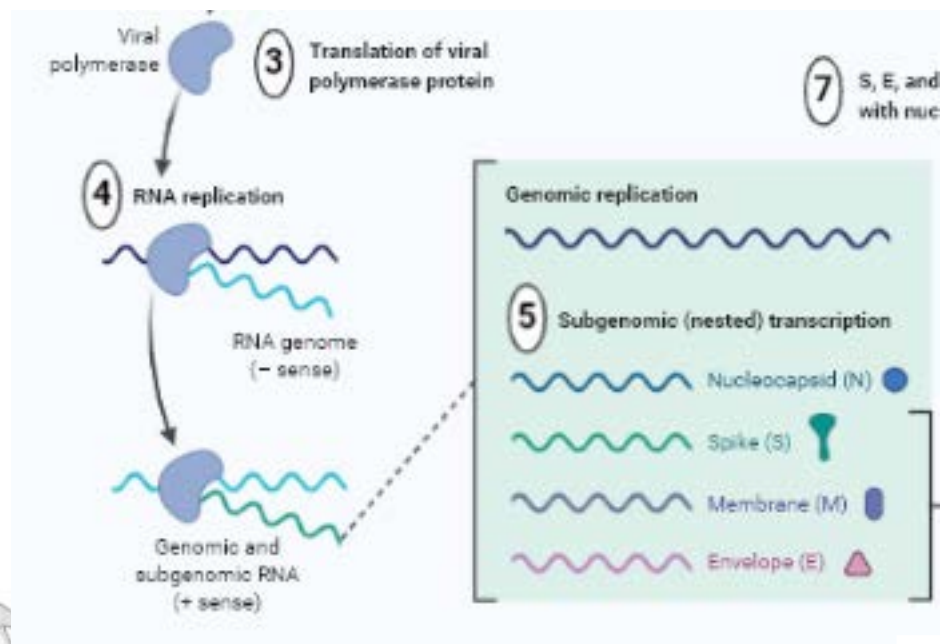
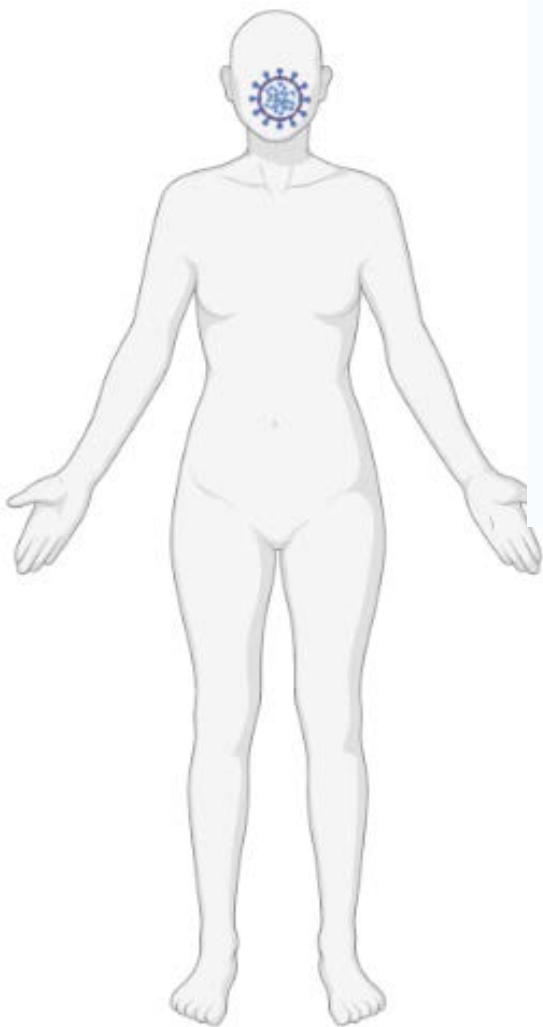
En menos de cuatro meses



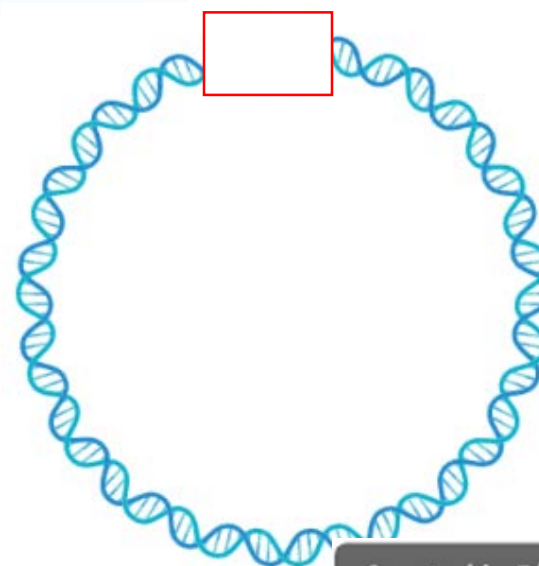
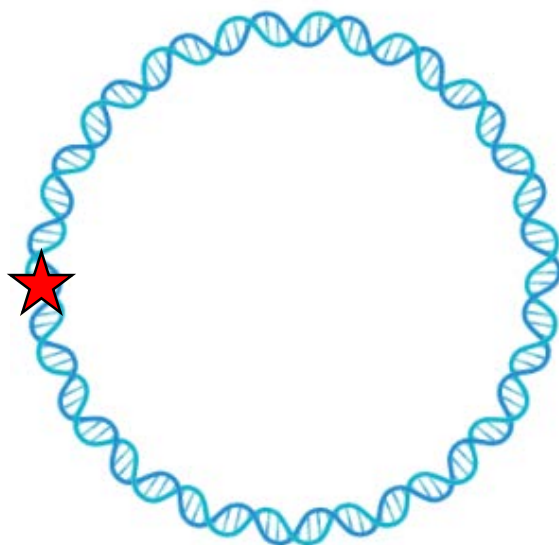
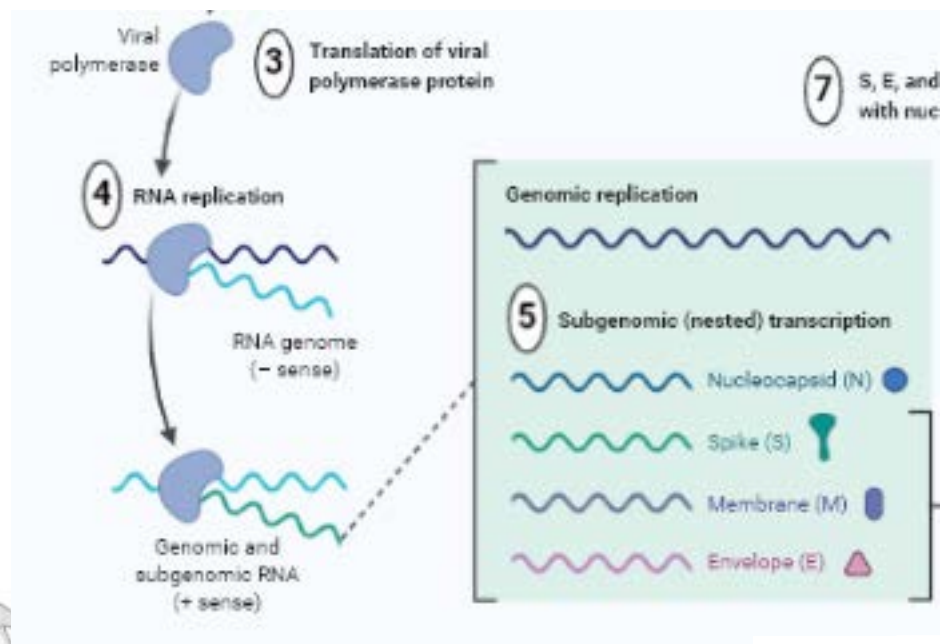
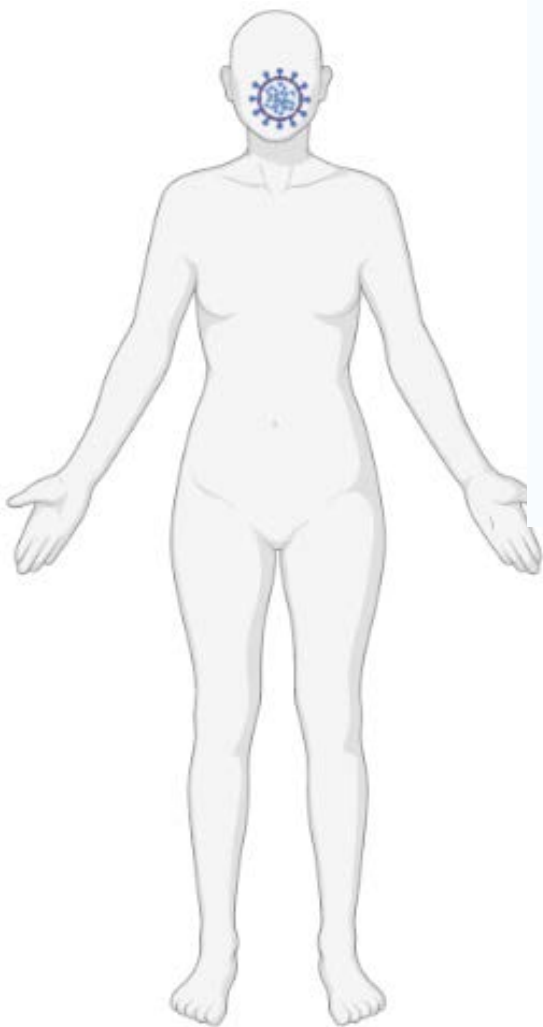
Variabilidad en los genomas de patógenos



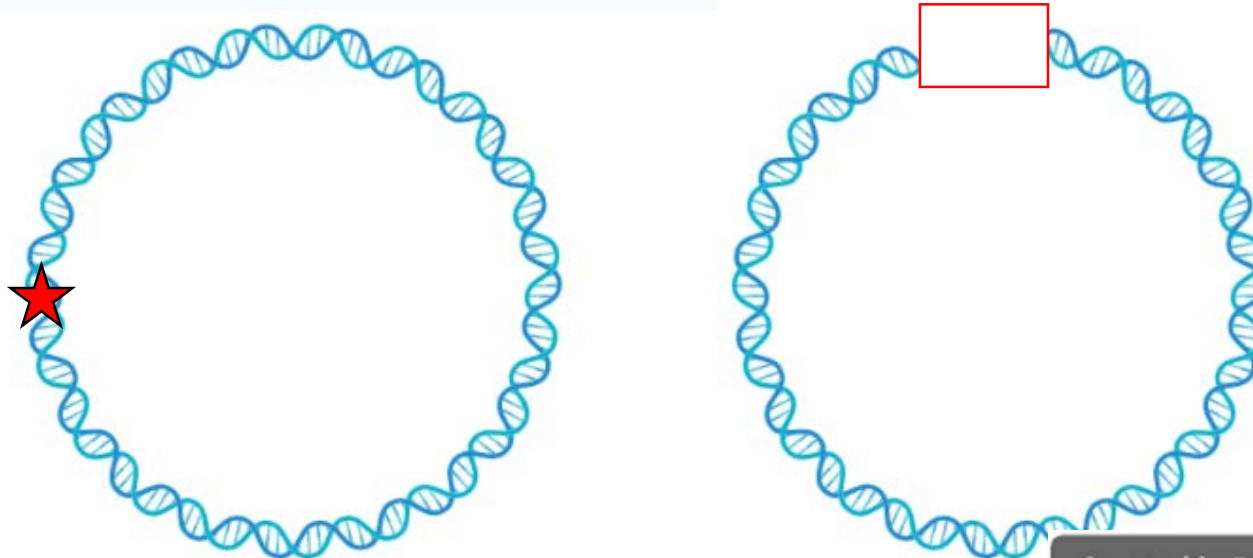
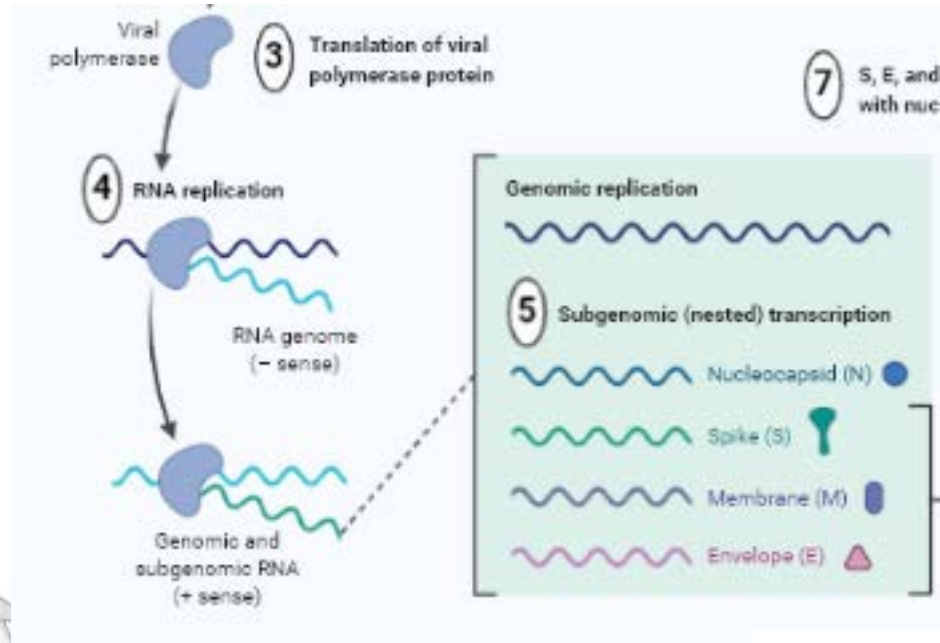
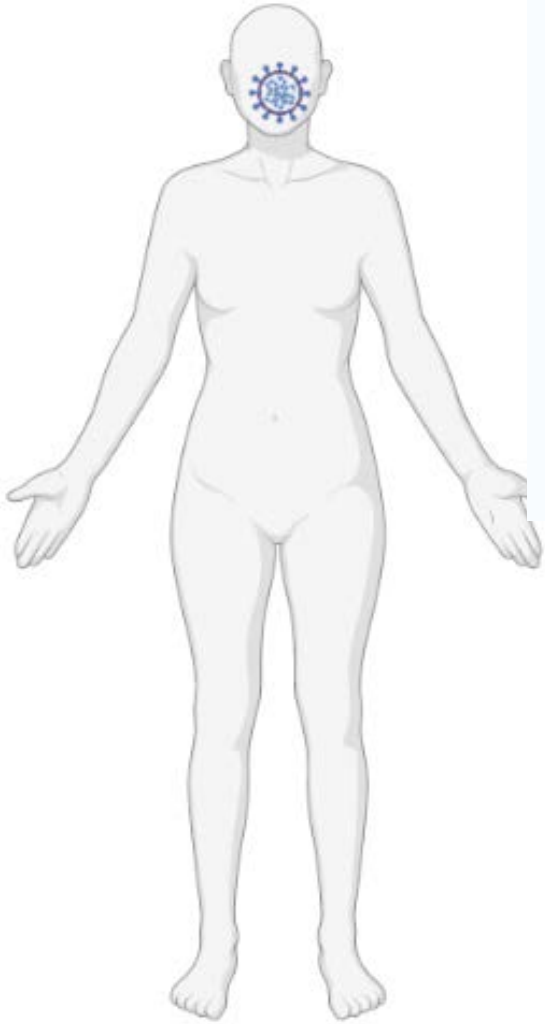
Variabilidad en los genomas de patógenos



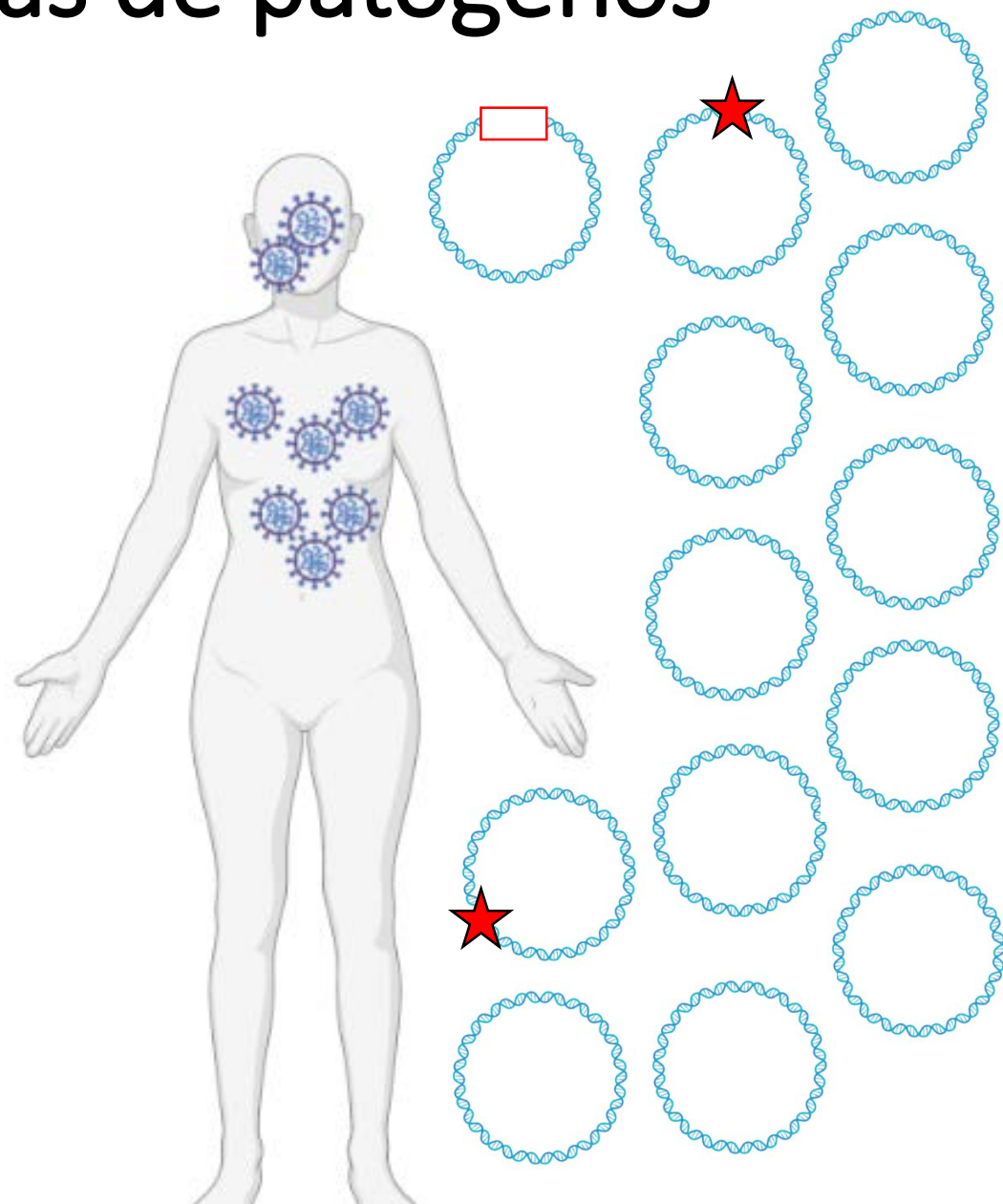
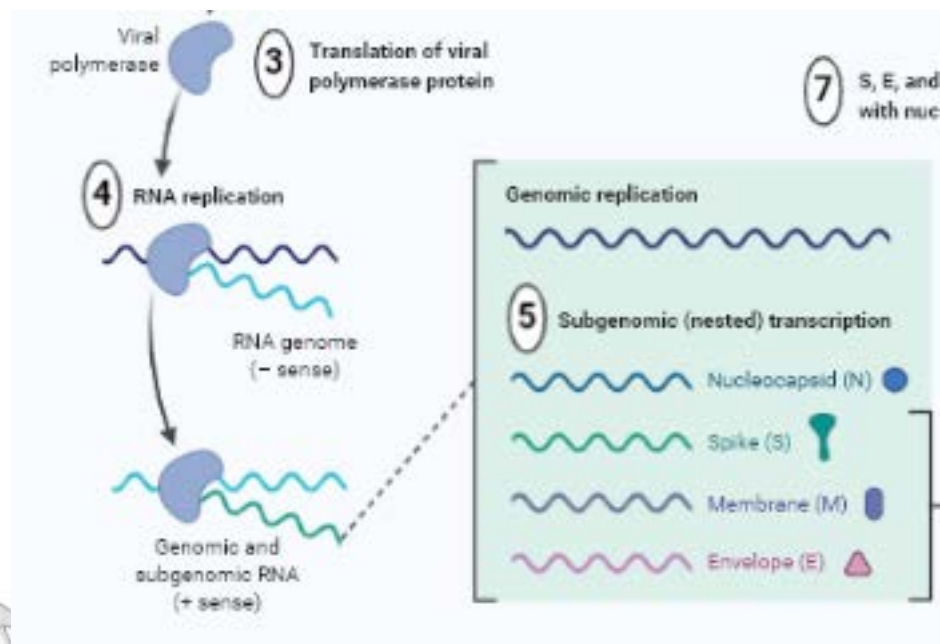
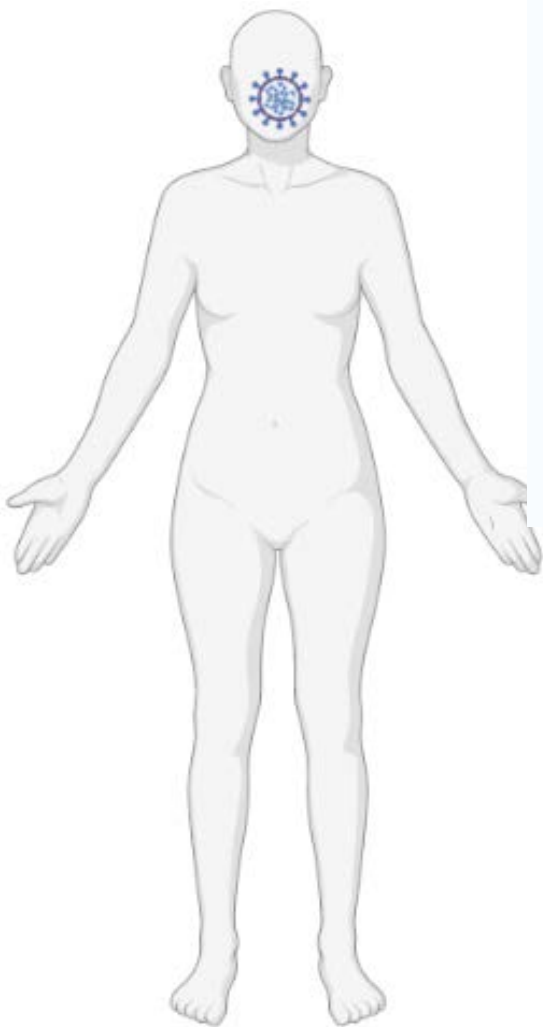
Variabilidad en los genomas de patógenos



Variabilidad en los genomas de patógenos



Variabilidad en los genomas de patógenos

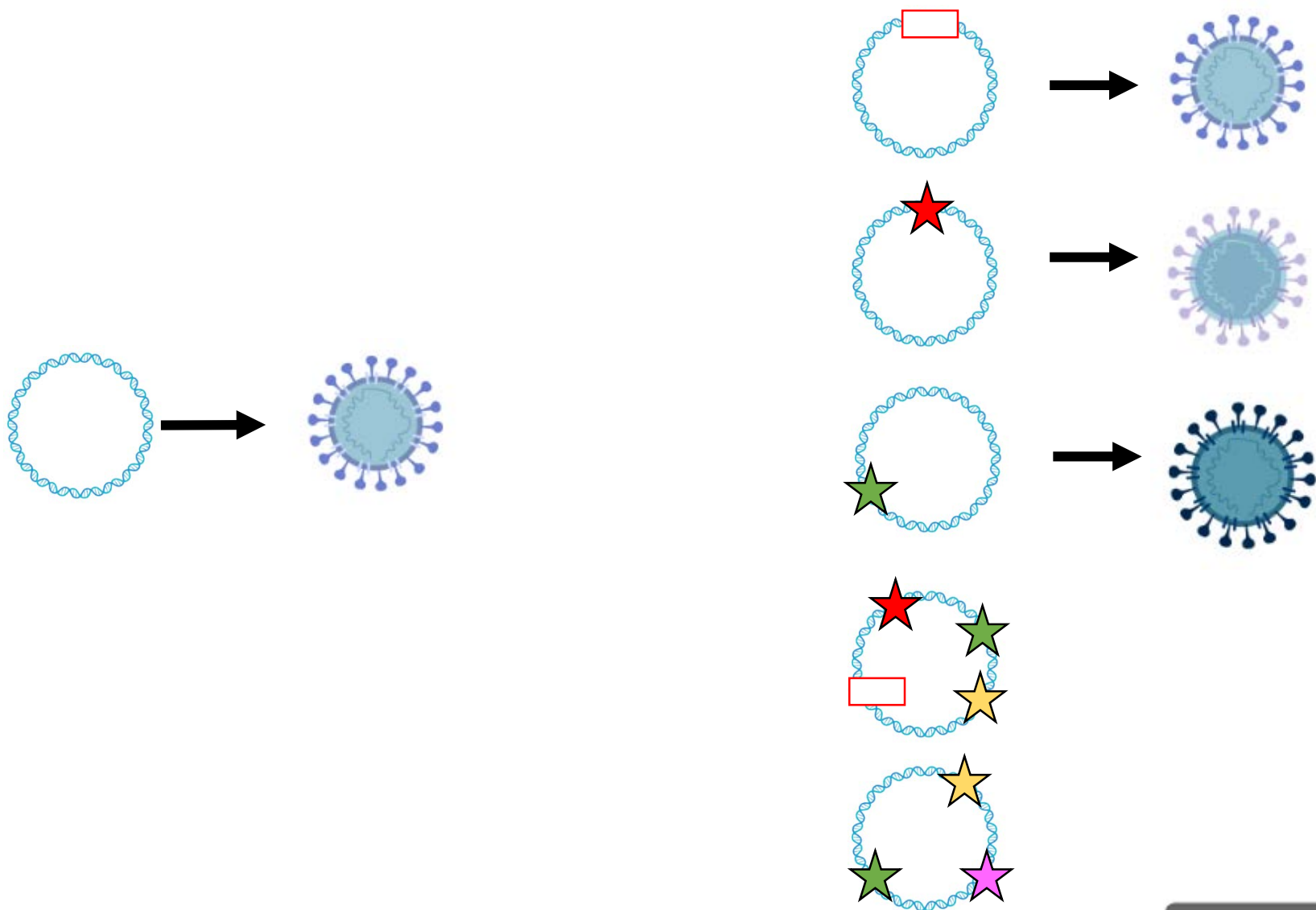


Variabilidad en los genomas

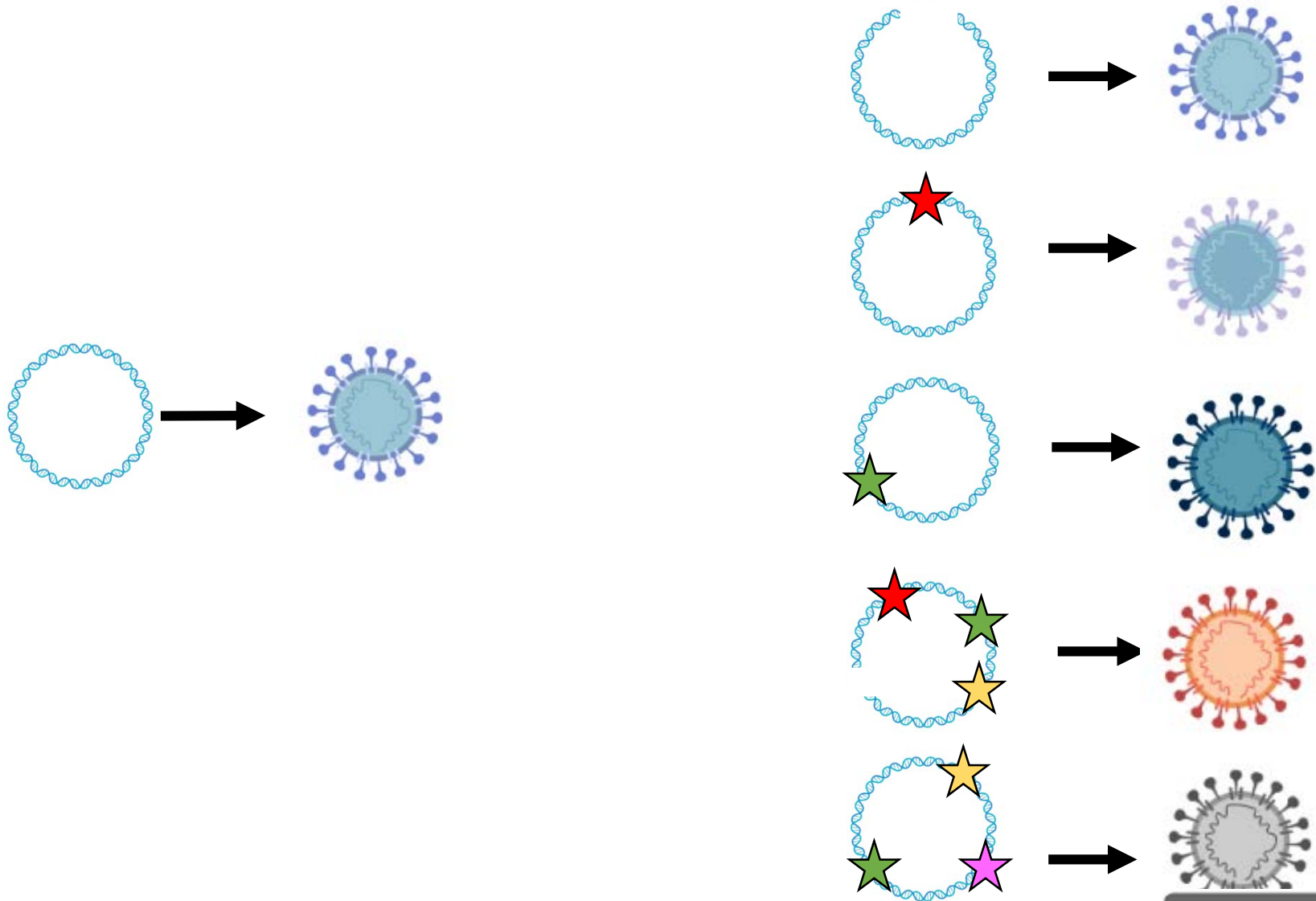
¿cuándo consideramos
que es una nueva cepa?

Cuando cambia sus características
biológicas

Variabilidad en los genomas de patógenos



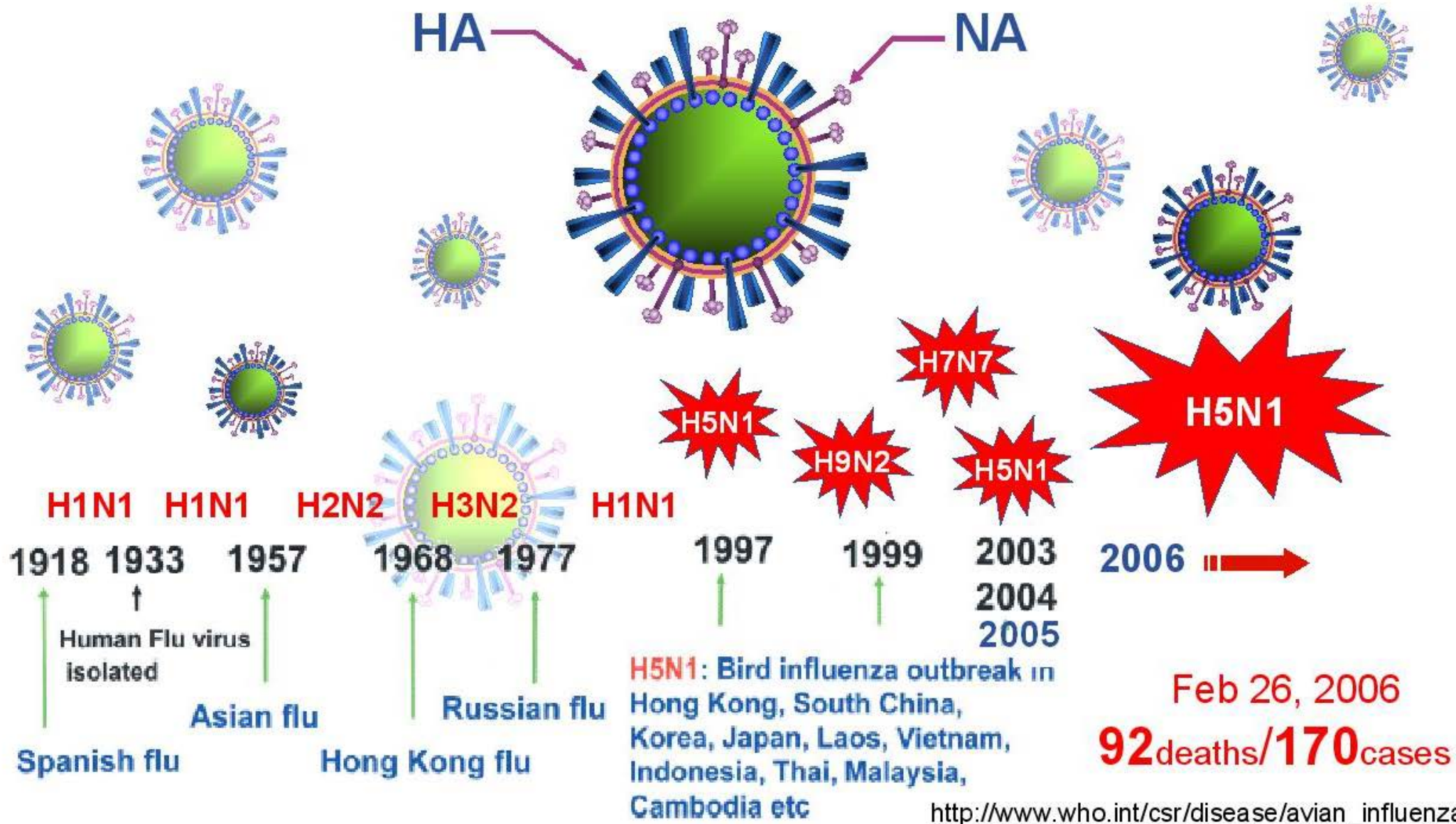
Variabilidad en los genomas de patógenos



El virus de la gripe es variable

H1, H2, H3, H4, ... H10

N1, N2, N3, N4, ... N7



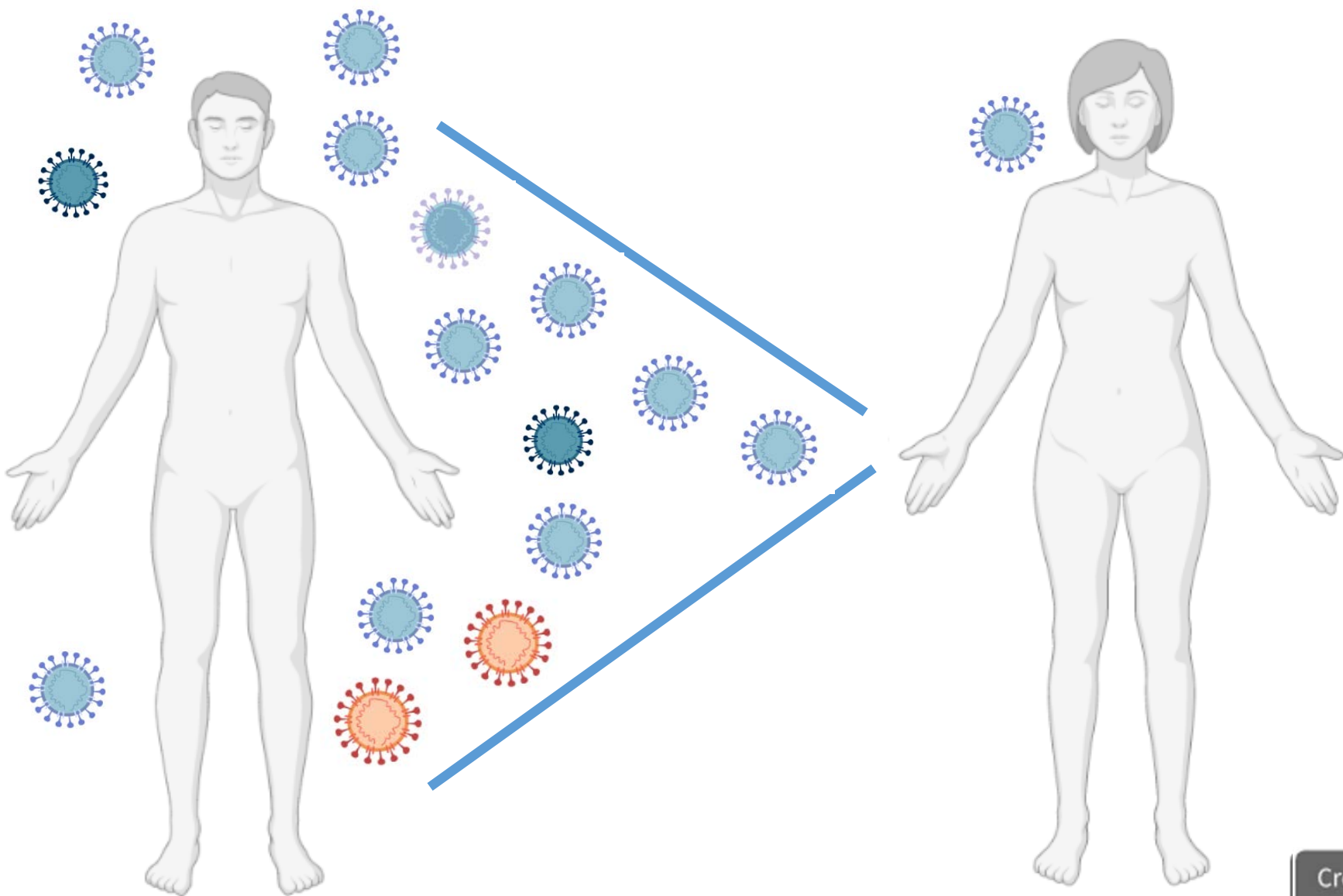
Variabilidad en los genomas de patógenos

¿cuándo consideramos que es una nueva cepa de un patógeno?

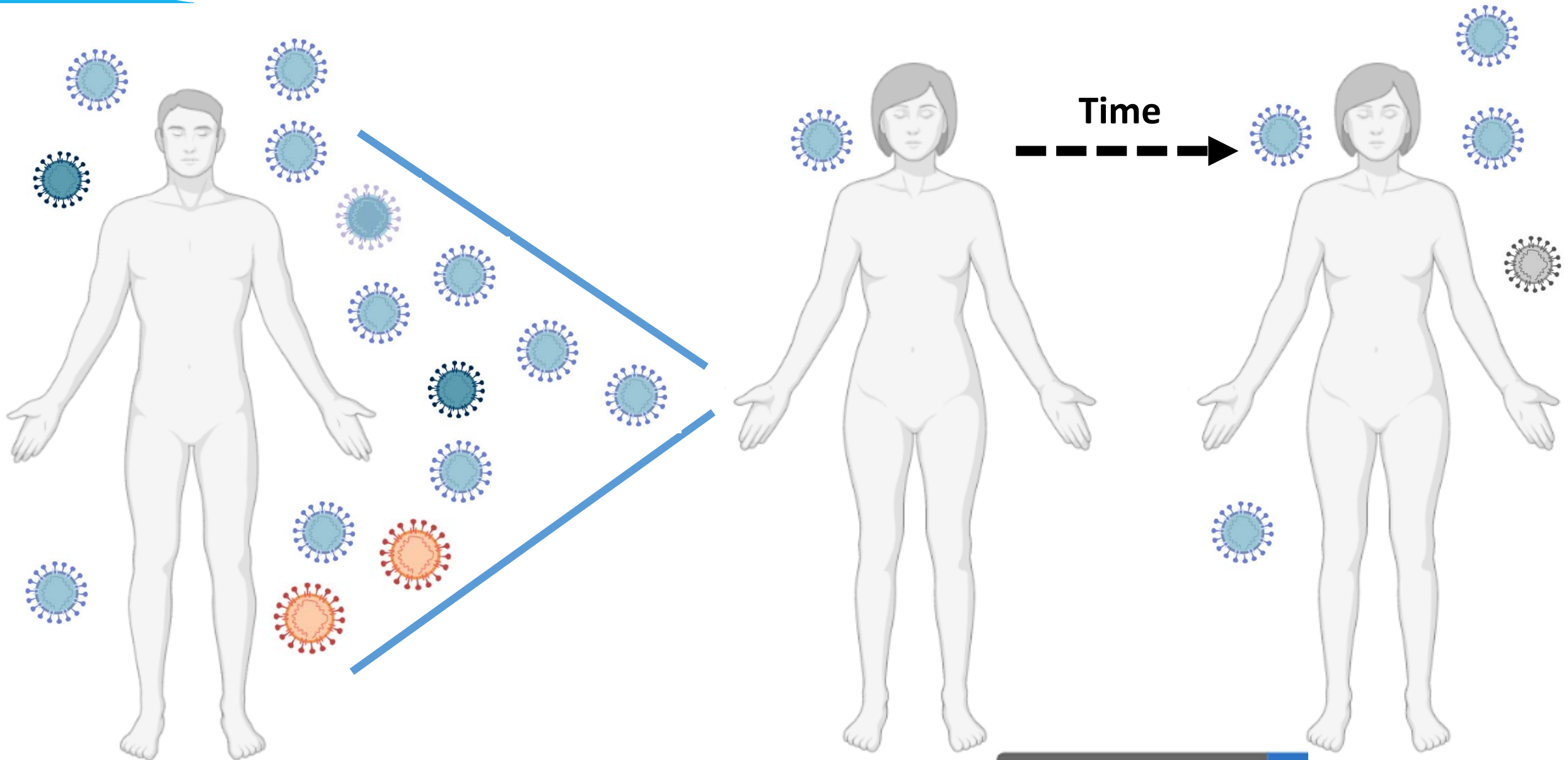
Frecuentemente, cuando las mutaciones producen un cambio en:

- la antigenicidad,**
- la especificidad de hospedador,**
- la clínica.**

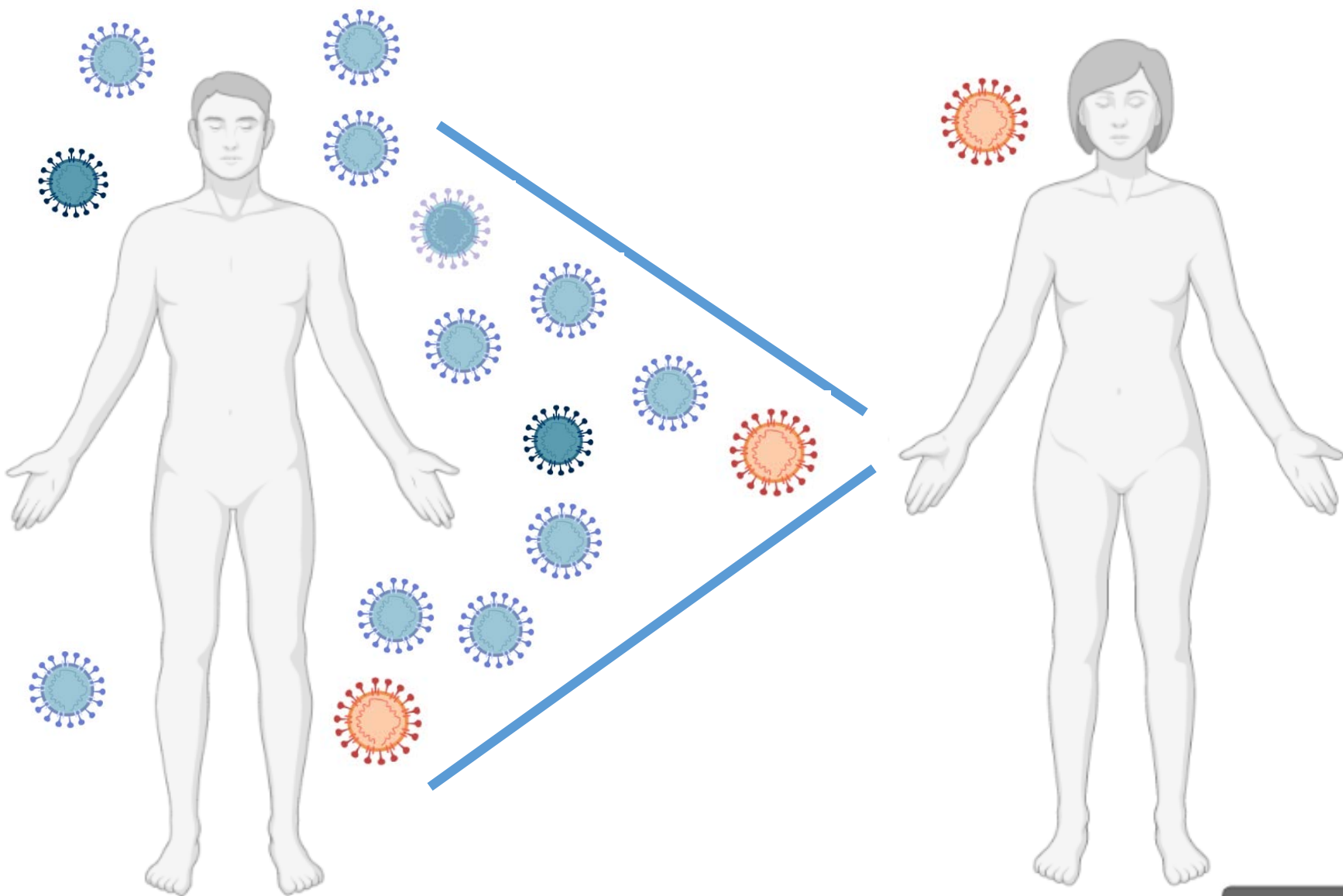
Variabilidad en los genomas de patógenos



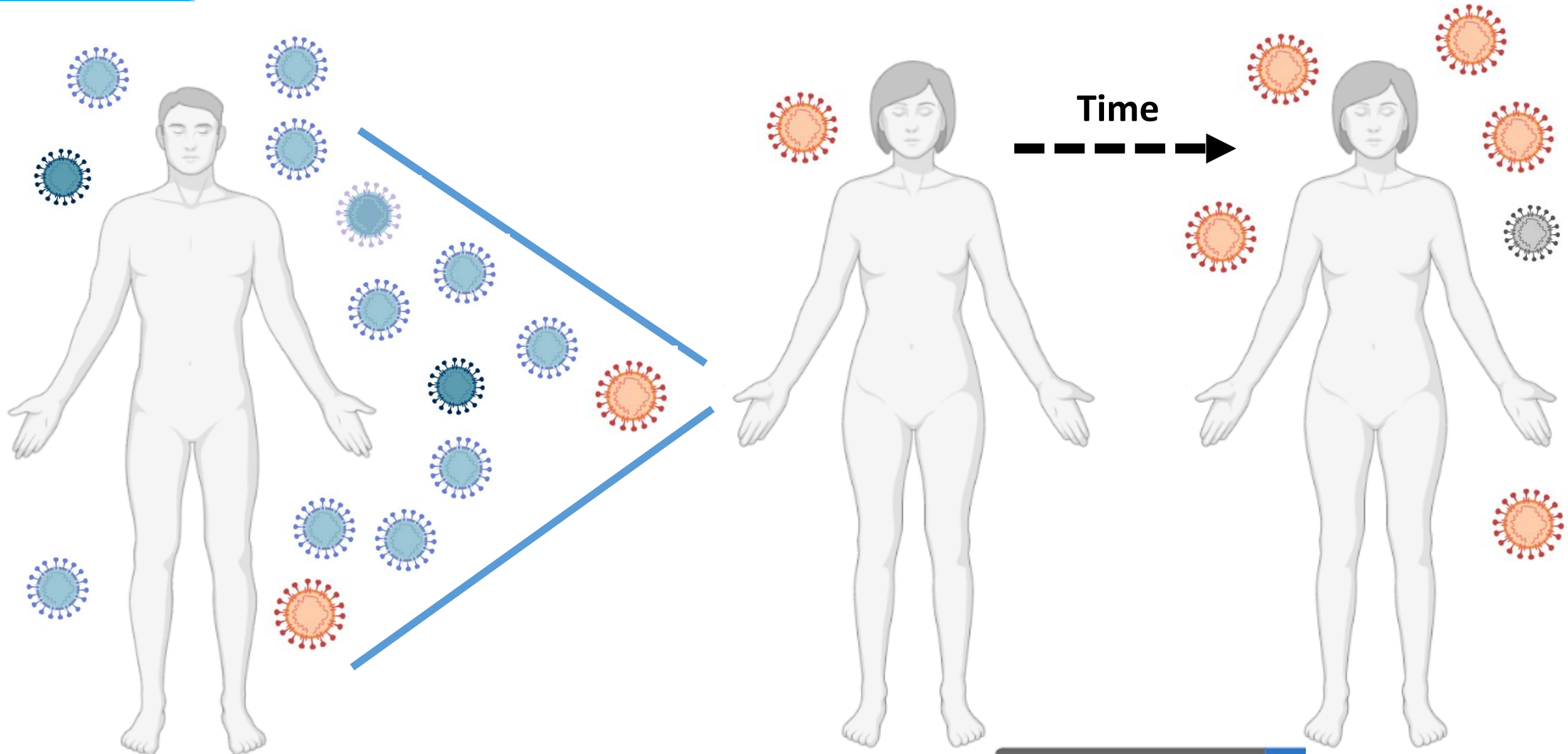
Variabilidad en los genomas de patógenos



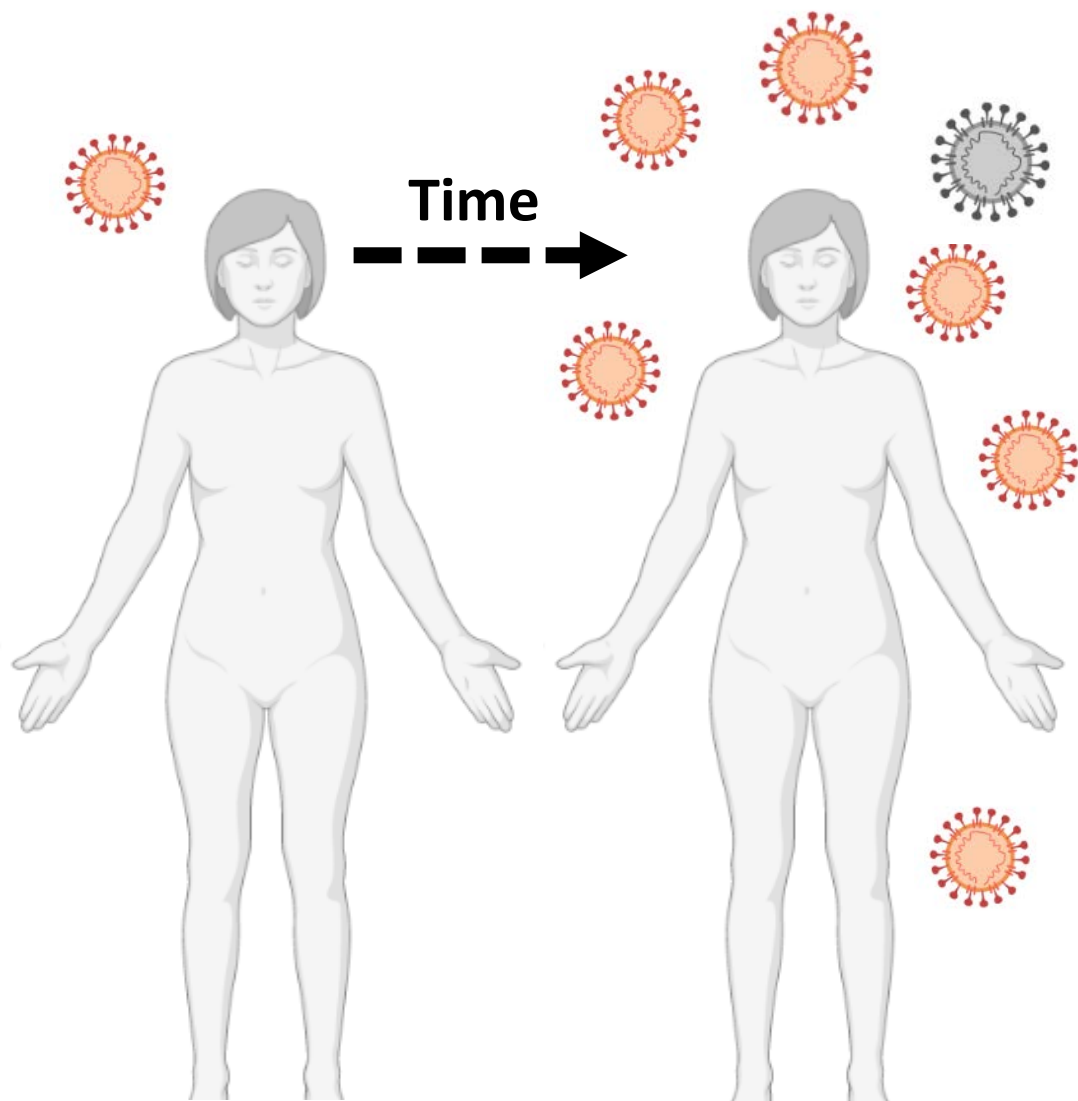
Variabilidad en los genomas de patógenos



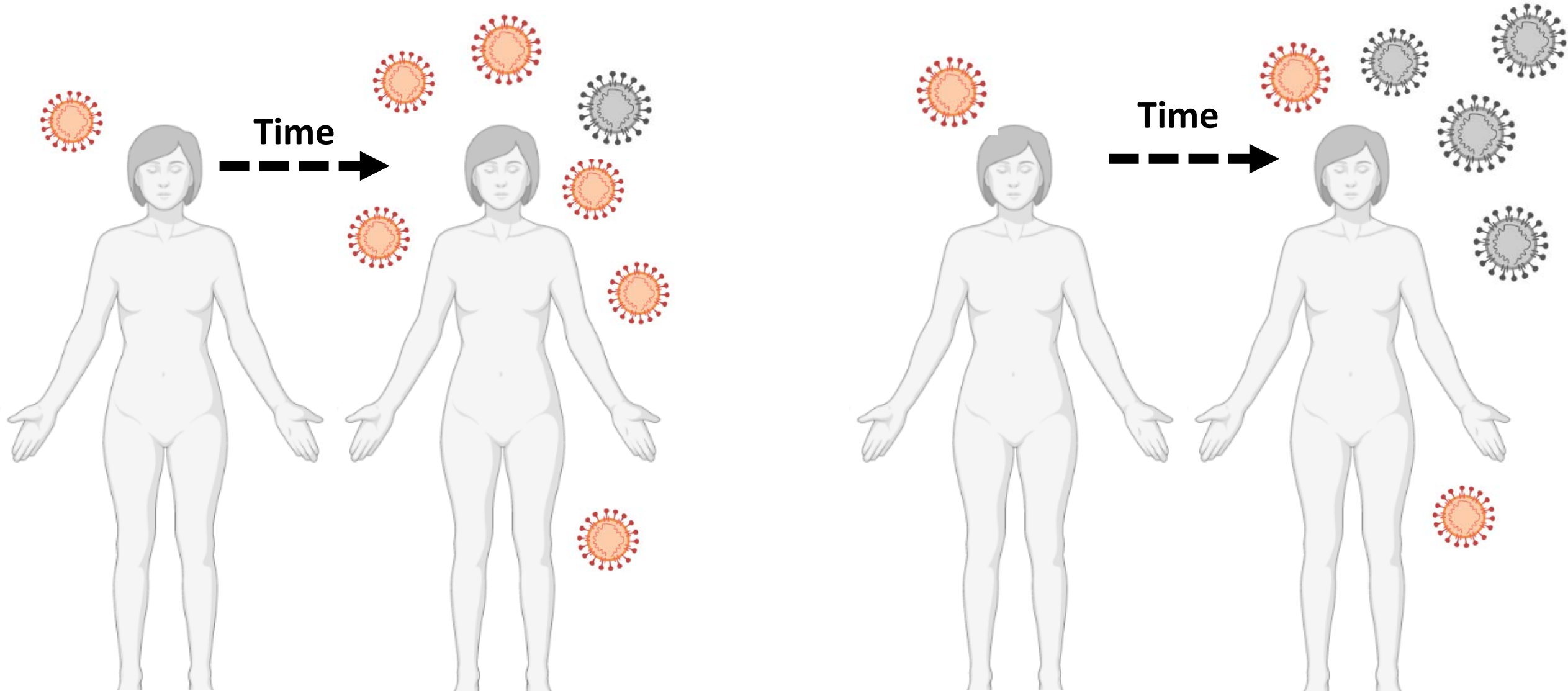
Variabilidad en los genomas de patógenos



Variabilidad en los genomas de patógenos



Variabilidad en los genomas de patógenos



Distribución de los 4,444,670 casos detectados



COVID-19 Dashboard by the
Center for Systems Science and
Engineering (CSSE) at Johns
Hopkins University (JHU)



Distribución de 25K secuencias del virus disponibles



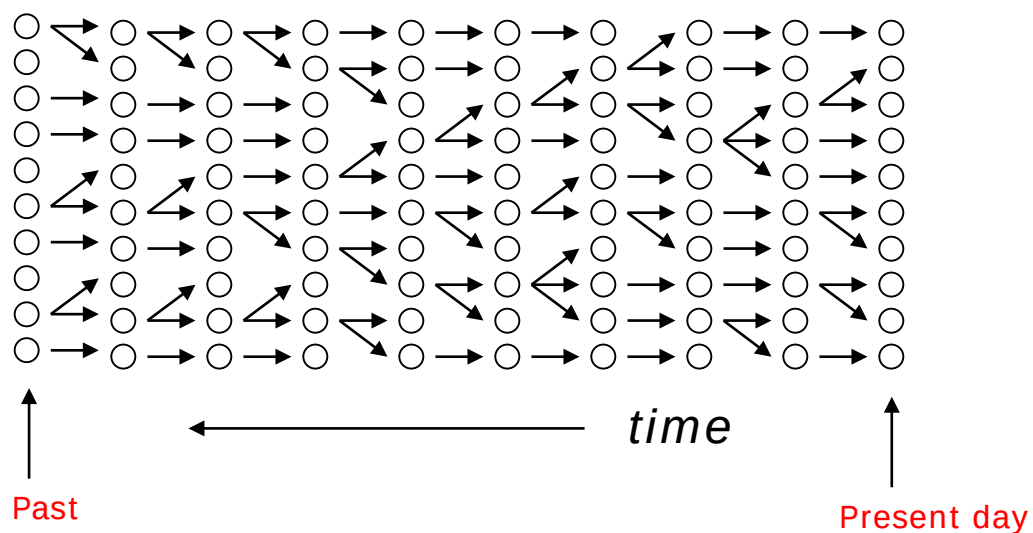
Nextstrain

Leaflet | © Mapbox © OpenStreetMap Improve this map

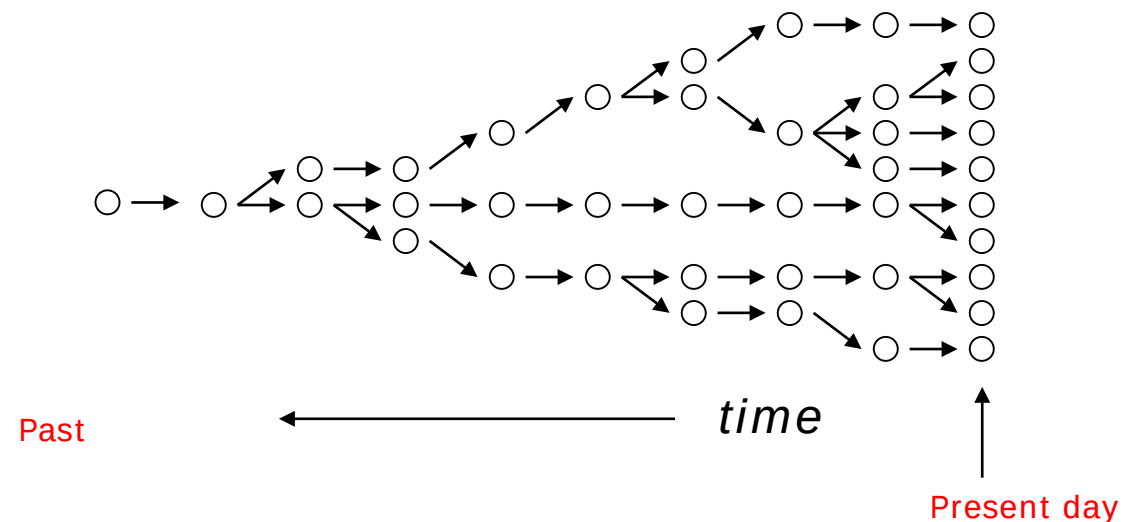
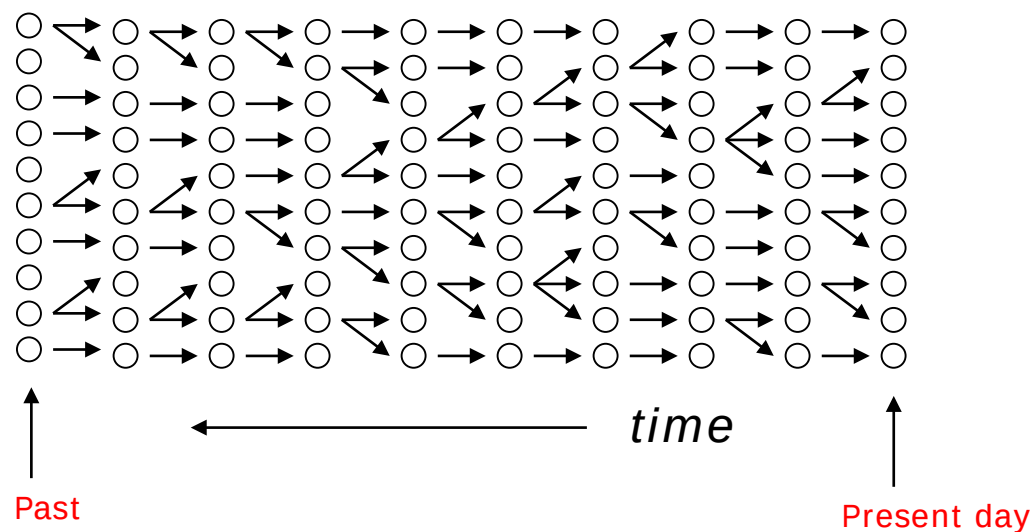
Genómica del SARS-CoV-2

- 1. Ampliar la comprensión de la biología del virus**
- 2. Evaluación y seguimiento de los diagnósticos, vacunas y tratamientos**
- 3. Mejorar nuestra comprensión de la epidemiología y la transmisión del SARS-CoV-2**

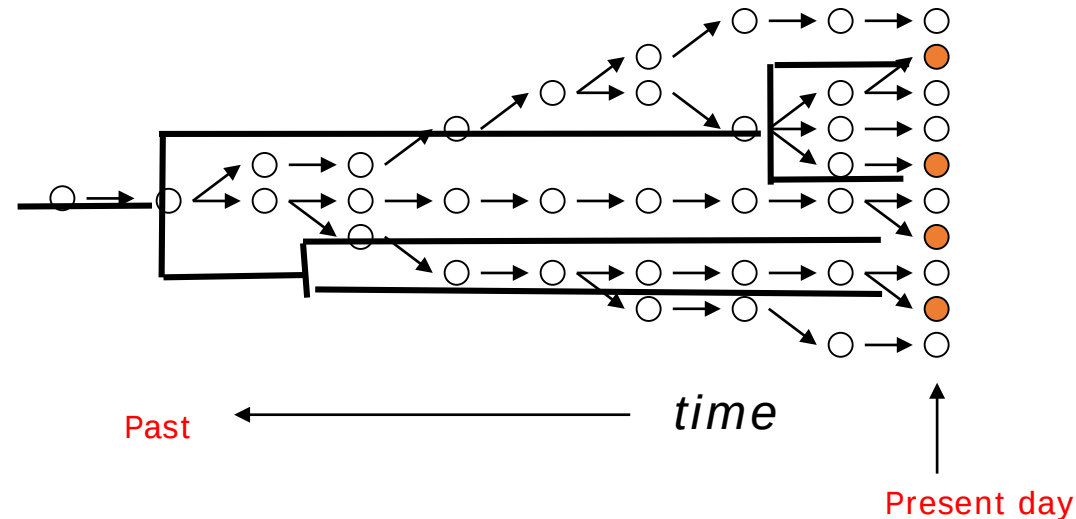
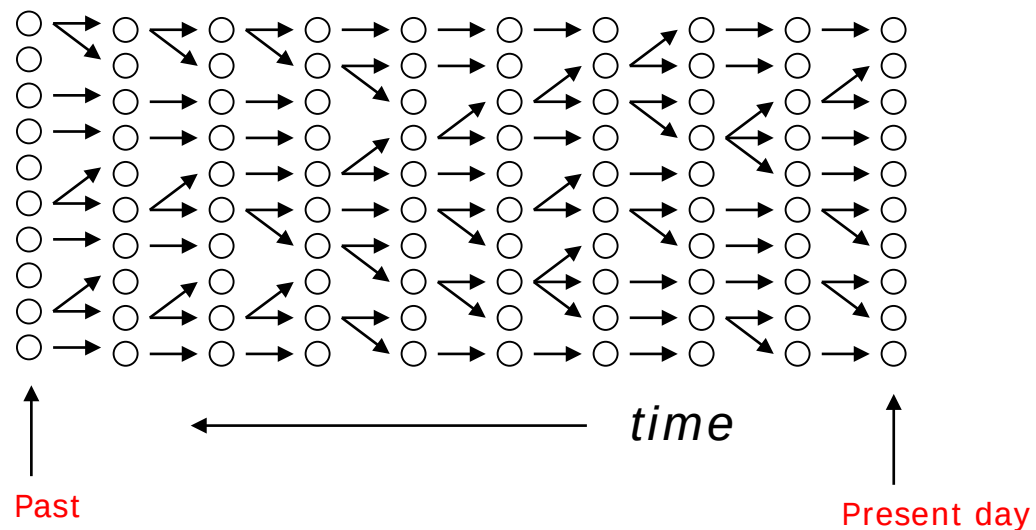
Filogenias: cómo reconstruir la historia



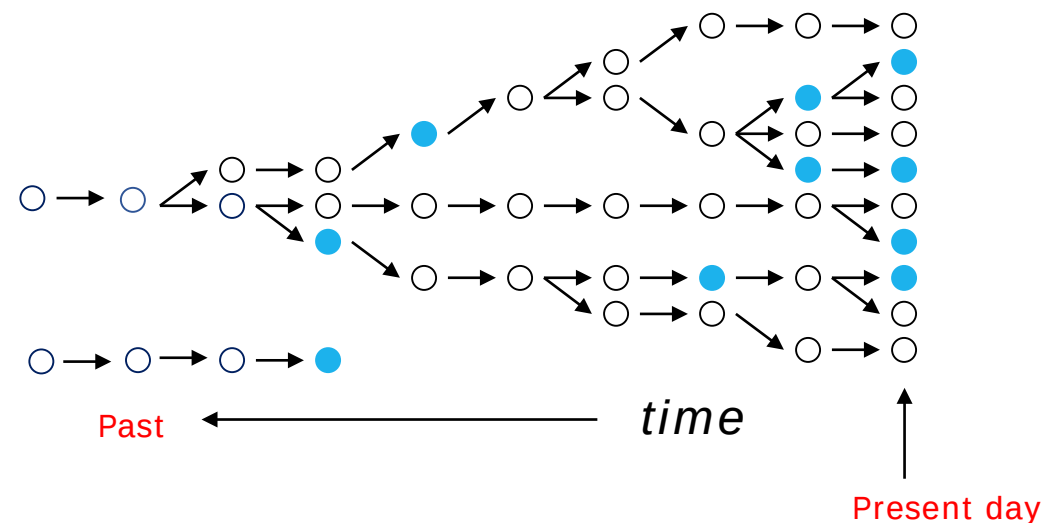
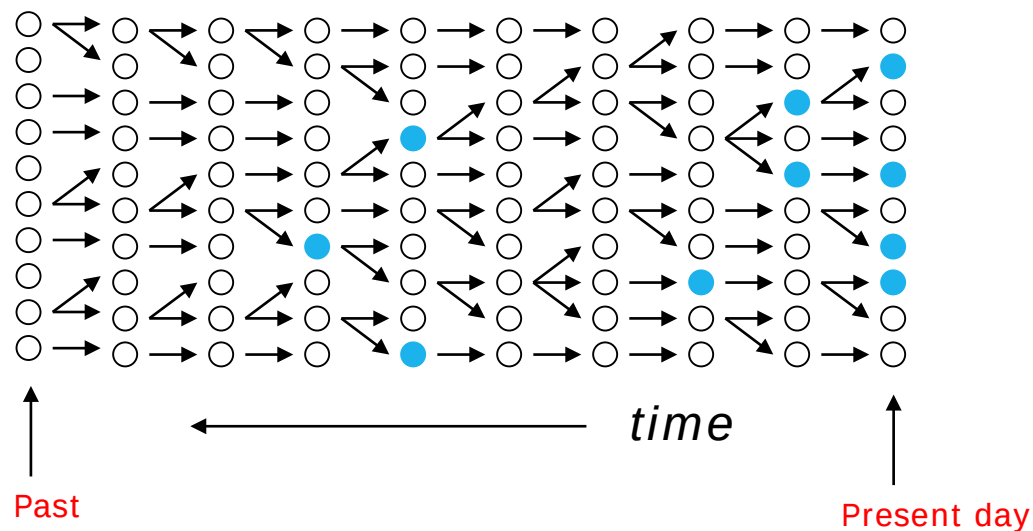
Filogenias: cómo reconstruir la historia



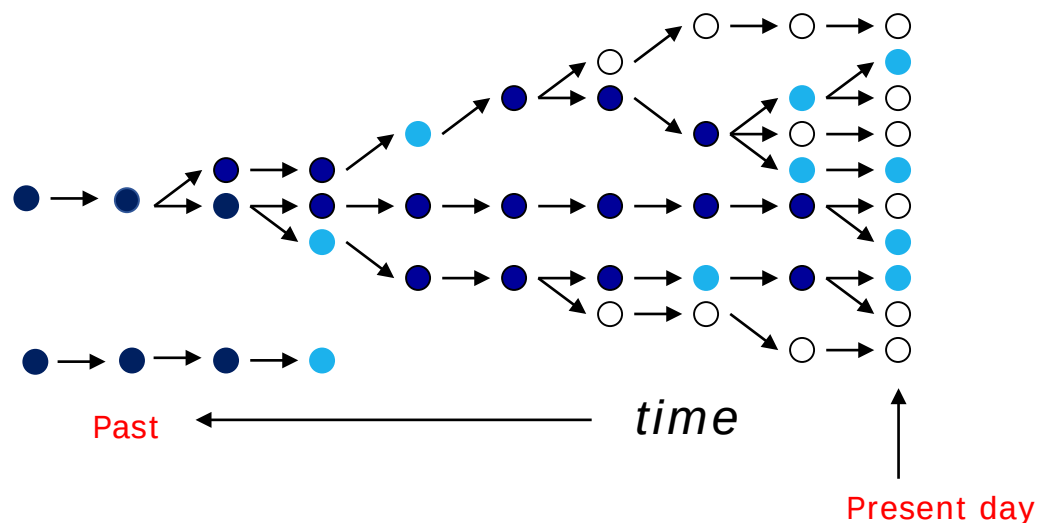
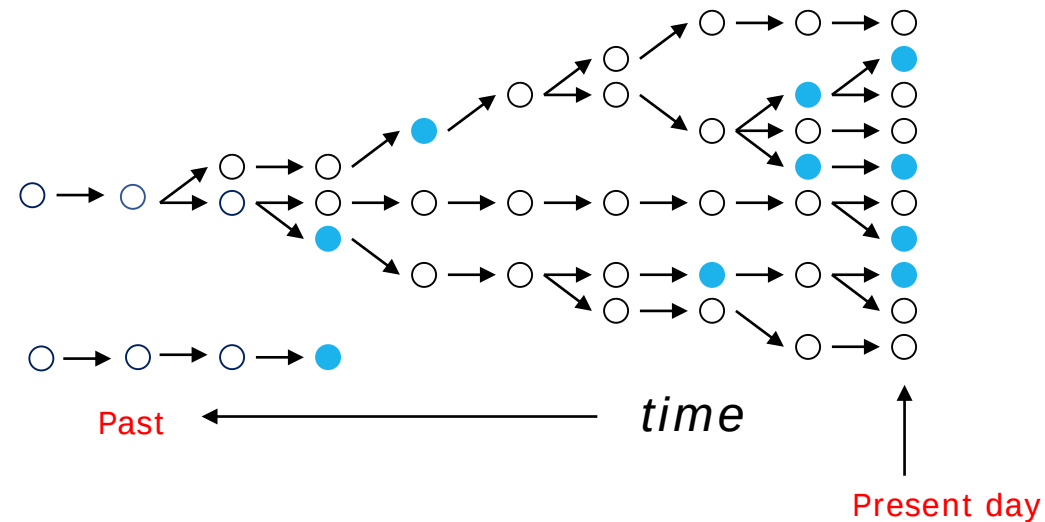
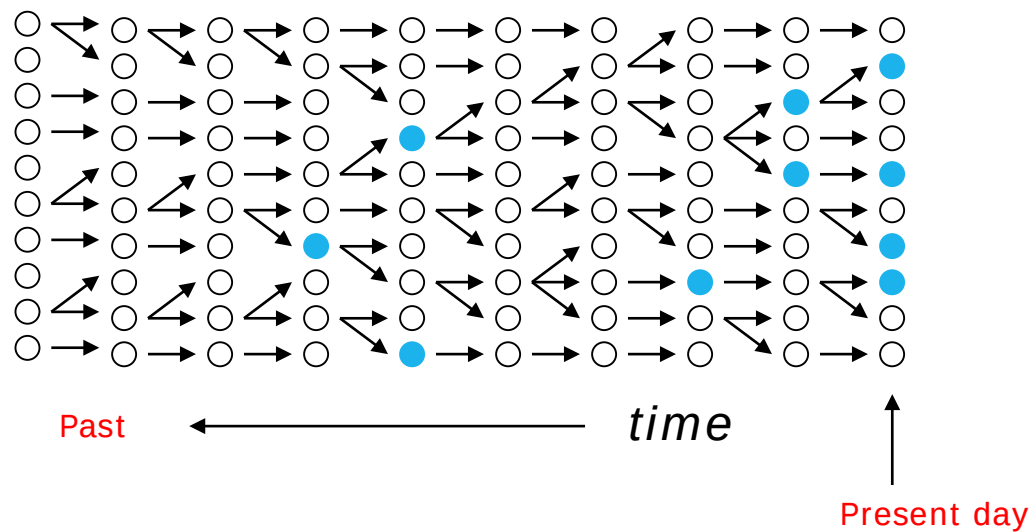
Filogenias: cómo reconstruir la historia



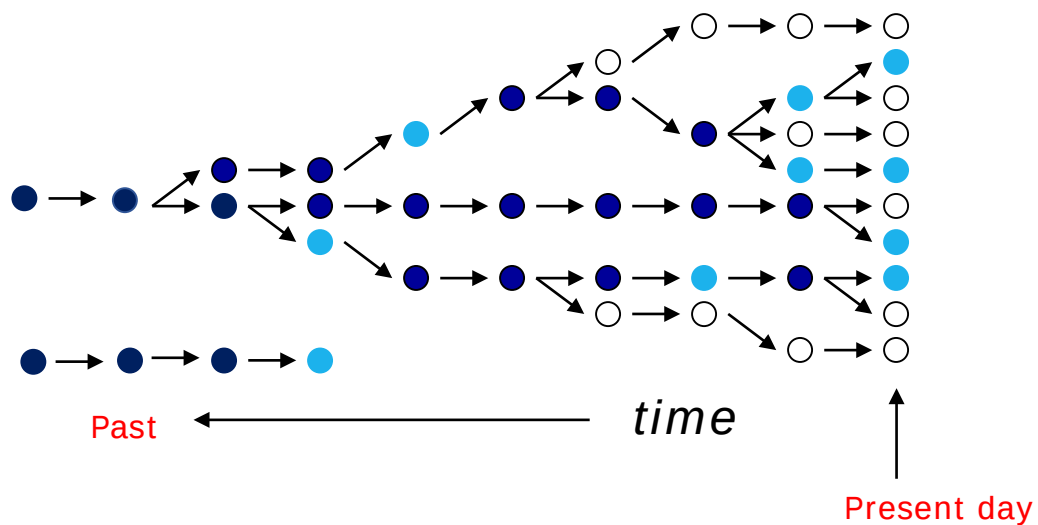
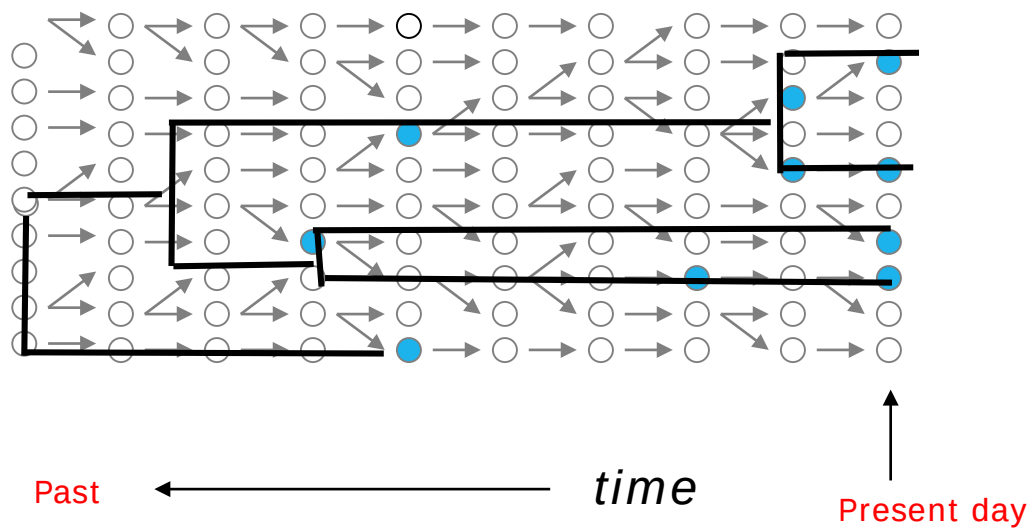
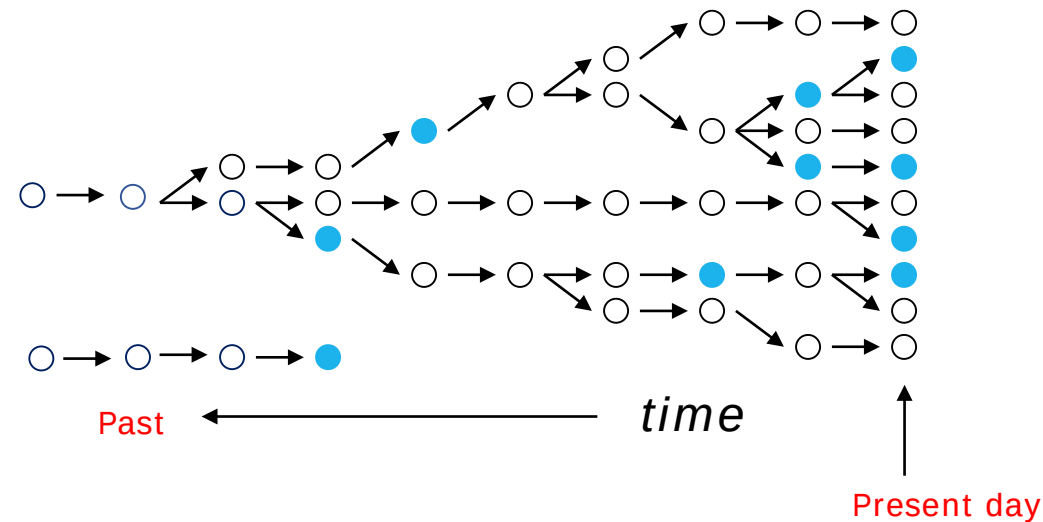
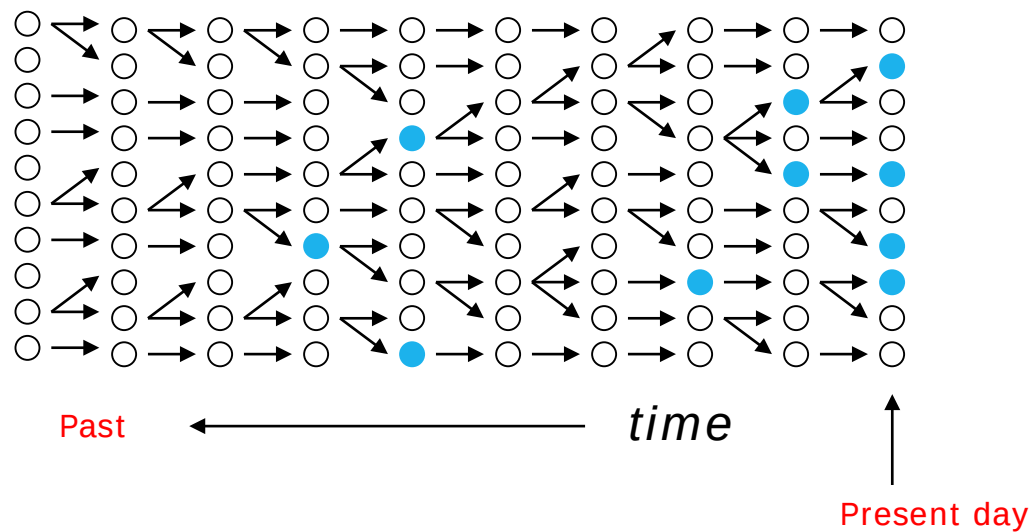
Filogenias: cómo reconstruir la historia



Filogenias: cómo reconstruir la historia



Filogenias: cómo reconstruir la historia



Epidemiología genómica

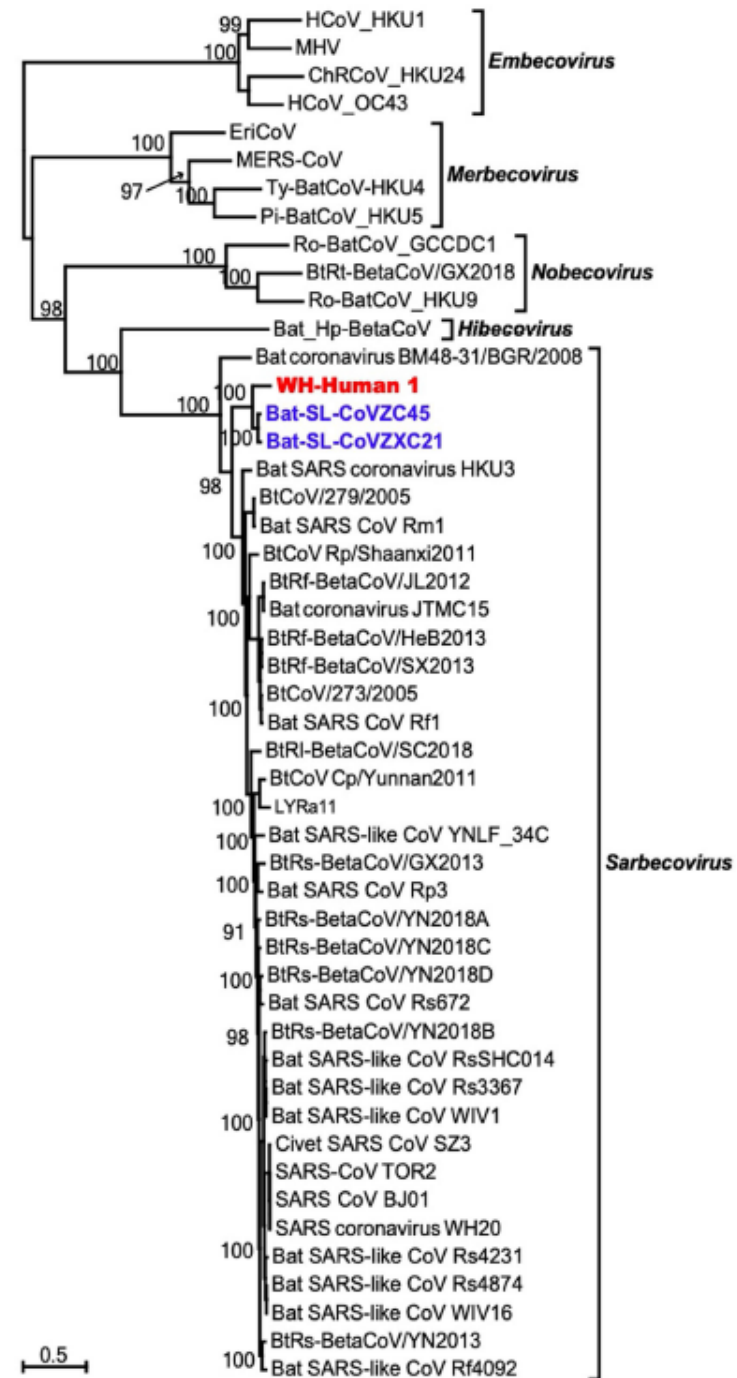
1. Ampliar la comprensión de la biología del virus

- Estudiar el origen del virus
- Evaluar la relevancia funcional y fenotípica de las mutaciones.

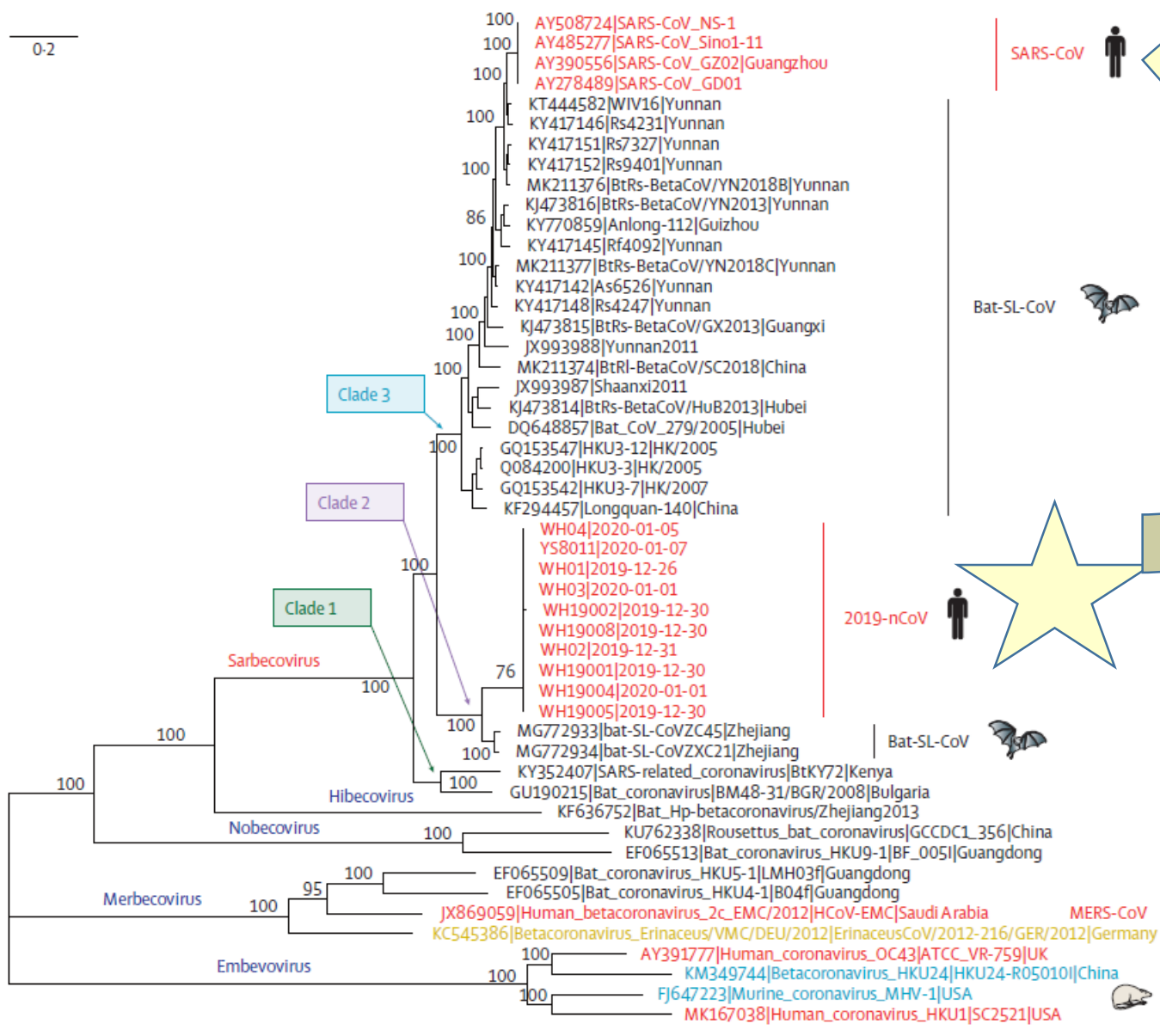
2. Permitir el seguimiento de los diagnósticos, vacunas y tratamientos.

3. Mejorar nuestra comprensión de la epidemiología y la transmisión del SARS-CoV-2.

¿a qué se parece?



¿a qué se parece?



El análisis filogenético ha revelado que el SARS-CoV-2 es un nuevo beta-CoV distinto de SARS-CoV y MERS-CoV

¿a qué se parece?

89%

RaTG13

Rhinolophus
affinis



84.7%

P5L/2017

Manis
javanica



97.2%

RmYN02

Rhinolophus
malayanus



Zhou et al.
doi.org/10.1016/j.cub.2020.05.023

Proteínas estructurales

ORF1ab

Spike

E M

N

Proteínas no estructurales (NSP)

Proteínas accesorias



¿a qué se parece?

RaTG13

92.9%

Rhinolophus
affinis



P5L/2017

83.2%

Manis
javanica



RmYN02

71.9%

Rhinolophus
malayanus



Zhou et al.
doi.org/10.1016/j.cub.2020.05.023

Proteínas estructurales

ORF1ab

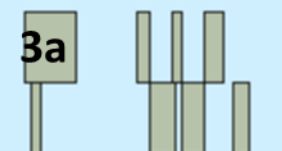
Spike

E M

N

Proteínas no estructurales (NSP)

Proteínas accesorias



¿a qué se parece?

RaTG13
92.9%

Rhinolophus
affinis



P5L/2017
85.2%

Manis
javanica



RmYN02
93.3%

Rhinolophus
malayanus



Zhou et al.
doi.org/10.1016/j.cub.2020.05.023

Proteínas estructurales

ORF1ab

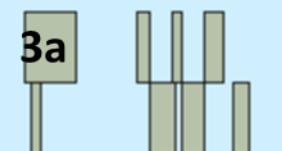
Spike

E M

N

Proteínas no estructurales (NSP)

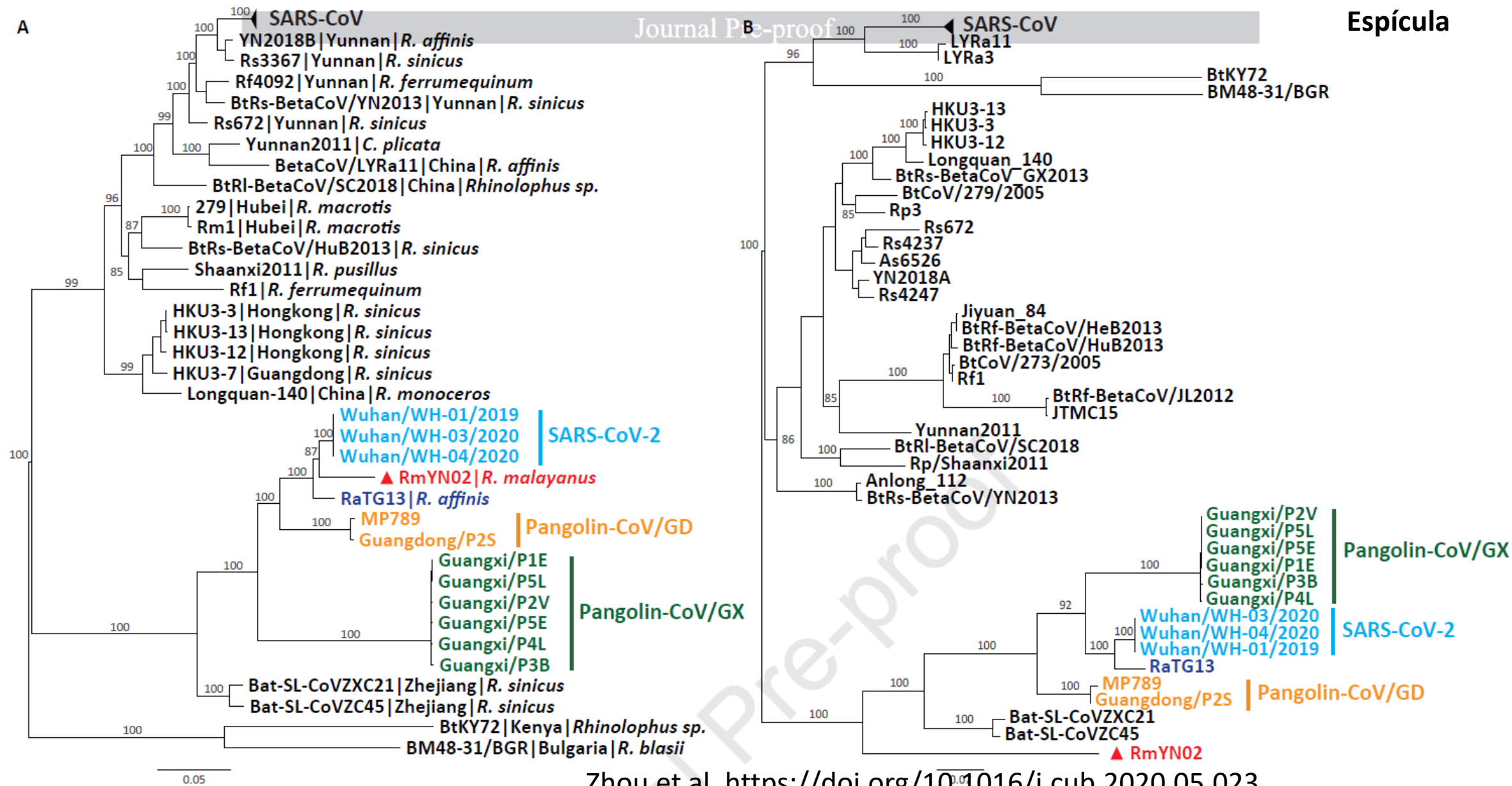
Proteínas accesorias



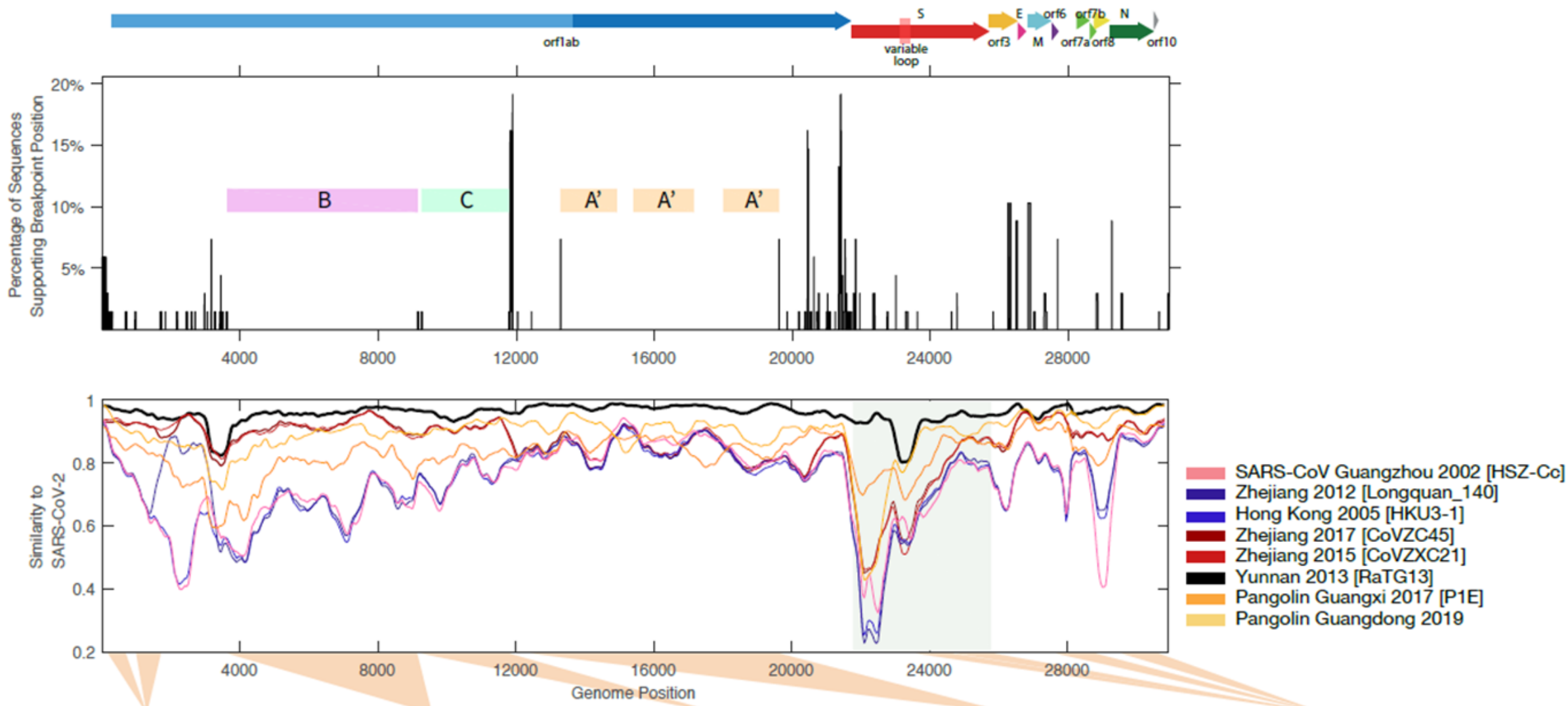
Cada región se parece a un virus

Genoma A

Espícula



Recombinación en la evolución del SARS-CoV-2



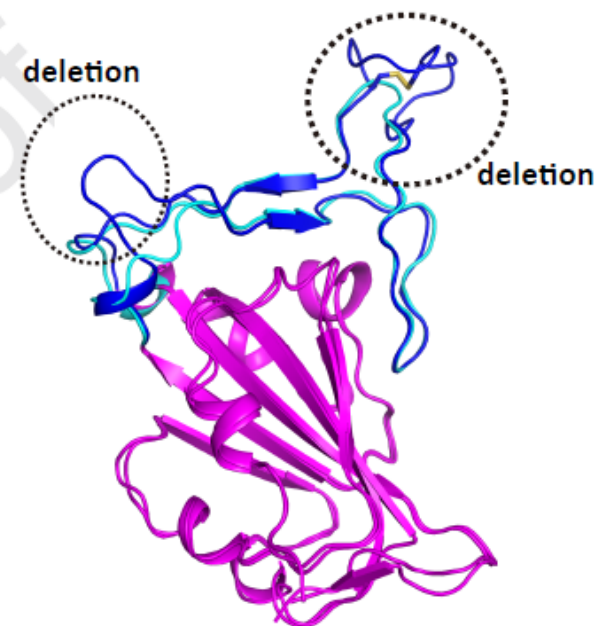
Insercion comun en espícula de pangolin

Consensus SARS-CoV-2	Y	L	Y	R	L	F	R	K	S	N	L	K	P	F	E	R	D	I	S	T	E	I	Y	Q	A	G	S	T	P	C	N	G	V	E	G	F	N	C	Y	F	P	L	Q	S	Y	G	F	Q	P	T	N	G	V	G	Y	Q	P	Y	R
RmYN02	Y	F	Y	R	S	H	R	A	V	K	L	K	P	F	E	R	D	L	S	S	D	E	-	-	-	-	-	-	-	N	G	V	R	-	-	-	-	-	T	L	S	T	Y	D	F	N	P	N	V	P	L	D	Y	Q	A	T	R		
RaTG13	Y	L	Y	R	L	F	R	K	A	N	L	K	P	F	E	R	D	I	S	T	E	I	Y	Q	A	G	S	K	P	C	N	G	Q	T	G	L	N	C	Y	Y	P	L	Y	R	Y	G	F	Y	P	T	D	G	V	G	H	Q	P	Y	R
ZC45	Y	F	Y	R	S	H	R	S	T	K	L	K	P	F	E	R	D	L	S	S	D	E	-	-	-	-	-	-	-	N	G	V	R	-	-	-	-	-	T	L	S	T	Y	D	F	N	P	N	V	P	L	E	Y	Q	A	T	R		
ZXC21	Y	F	Y	R	S	H	R	S	T	K	L	K	P	F	E	R	D	L	S	S	D	E	-	-	-	-	-	-	-	N	G	V	R	-	-	-	-	-	T	L	S	T	Y	D	F	N	P	N	V	P	L	E	Y	Q	A	T	R		
pangolin/MP789/2019	Y	L	Y	R	L	F	R	K	S	N	L	K	P	F	E	R	D	I	S	T	E	I	Y	Q	A	G	S	T	P	C	N	G	V	E	G	F	N	C	Y	F	P	L	Q	S	Y	G	F	H	P	T	N	G	V	G	Y	Q	P	Y	R
pangolin/GX/P5L/2017	Y	L	Y	R	L	F	R	K	S	K	L	K	P	F	E	R	D	I	S	T	E	I	Y	Q	A	G	S	T	P	C	N	G	Q	V	G	L	N	C	Y	Y	P	L	E	R	Y	G	F	H	P	T	T	G	V	N	Y	Q	P	F	R
SARS-CoV GZ02	Y	K	Y	R	Y	L	R	H	G	K	L	R	P	F	E	R	D	I	S	N	V	P	F	S	P	D	G	K	P	C	T	P	-	P	A	L	N	C	Y	W	P	L	N	D	Y	G	F	Y	T	T	G	I	G	Y	Q	P	Y	R	
RmYN01	Y	Y	Y	R	S	S	R	K	T	K	L	K	P	F	E	R	D	L	S	S	D	E	-	-	-	-	-	-	-	N	G	V	R	-	-	-	-	-	T	L	S	T	Y	D	F	Y	P	S	V	P	L	E	Y	Q	A	T	R		
SARS-CoV-2 Numbering	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509

E SARS-CoV-2 / pangolin/MP789/2019



F SARS-CoV-2 / RmYN02



Origen evolutivo virus vs origen de epidemia

Un virus parecido a coronavirus de murciélagos



Origen virus vs epidemia



¿Mercado pescado Wuhan?

Necesitaremos tener muestras de los animals del mercado o cualquier otra possible fuente para secuenciar y comparar

Origen natural, no ingeniería genética



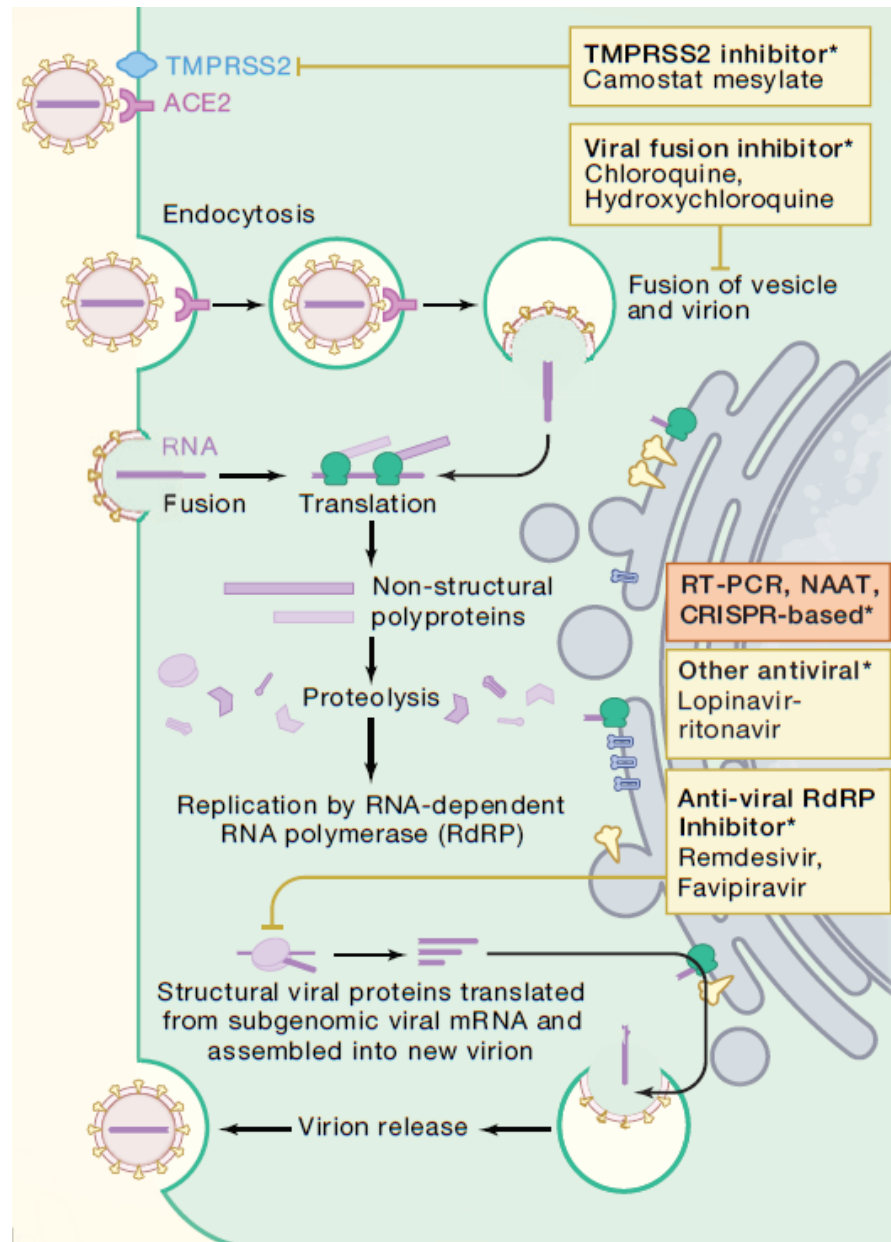
correspondence

The proximal origin of SARS-CoV-2

1. Selección artificial por ingeniería genética
2. Selección natural
 - 2.1 En animales no humanos y luego transmisión zoonótica a humanos
 - 2.2 En humanos y luego transferencia zoonótica

La secuencia de la espícula del SARS-CoV-2 no es la que los modelos predicen como la solución óptima en la unión con el receptor humano ACE2. Es más plausible que sea resultado de la evolución en humanos o en otro animal con un receptor similar

Mutaciones en proteínas virales

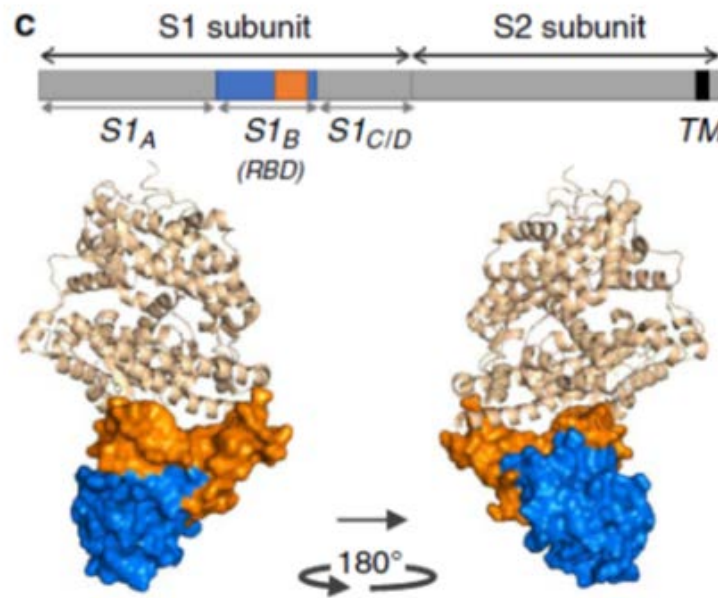


Algunos cambios en las proteínas estructurales **podrían** alterar la afinidad con las células huésped

Mutaciones en dianas de tratamientos

A human monoclonal antibody blocking SARS-CoV-2 infection

Chunyan Wang^{1,6}, Wentao Li^{1,6}, Dubravka Drabek^{2,3,6}, Nisreen M. A. Okba⁴, Rien van Haperen^{2,3}, Albert D. M. E. Osterhaus⁵, Frank J. M. van Kuppeveld¹, Bart L. Haagmans⁴, Frank Grosveld^{2,3,7} & Berend-Jan Bosch^{1,7}✉



Pachetti et al. *J Transl Med* (2020) 18:179
<https://doi.org/10.1186/s12967-020-02344-6>

Journal of
Translational Medicine

RESEARCH

Open Access



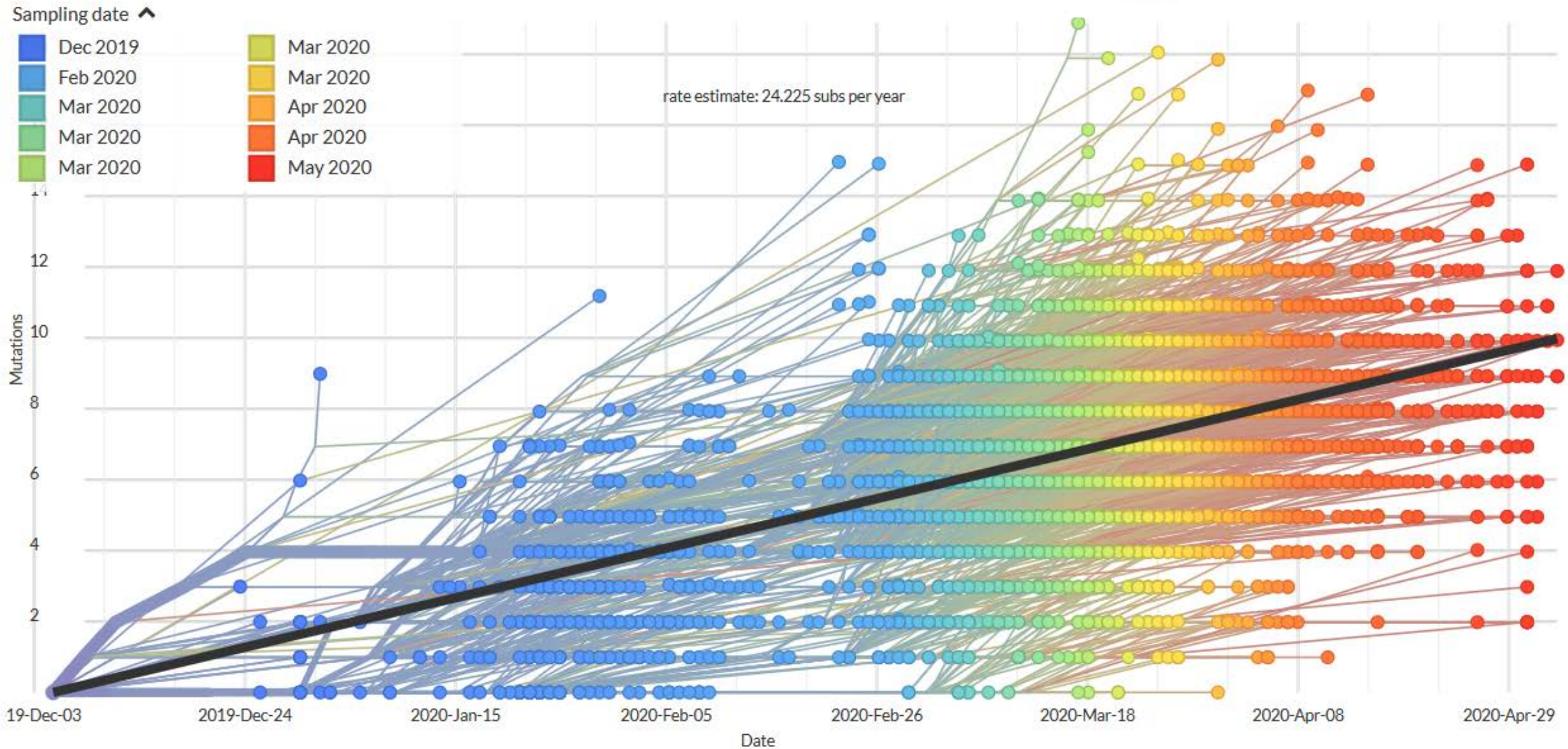
Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant

Maria Pachetti^{1,2†}, Bruna Marini^{4†}, Francesca Benedetti⁵, Fabiola Giudici³, Elisabetta Mauro⁴, Paola Storici¹, Claudio Masciovecchio¹, Silvia Angeletti⁶, Massimo Ciccozzi⁶, Robert C. Gallo^{7,8}, Davide Zella^{5,9*} and Rudy Ippodrino^{4*}

Acumulación de mutaciones



Nextstrain



Epidemiología genómica

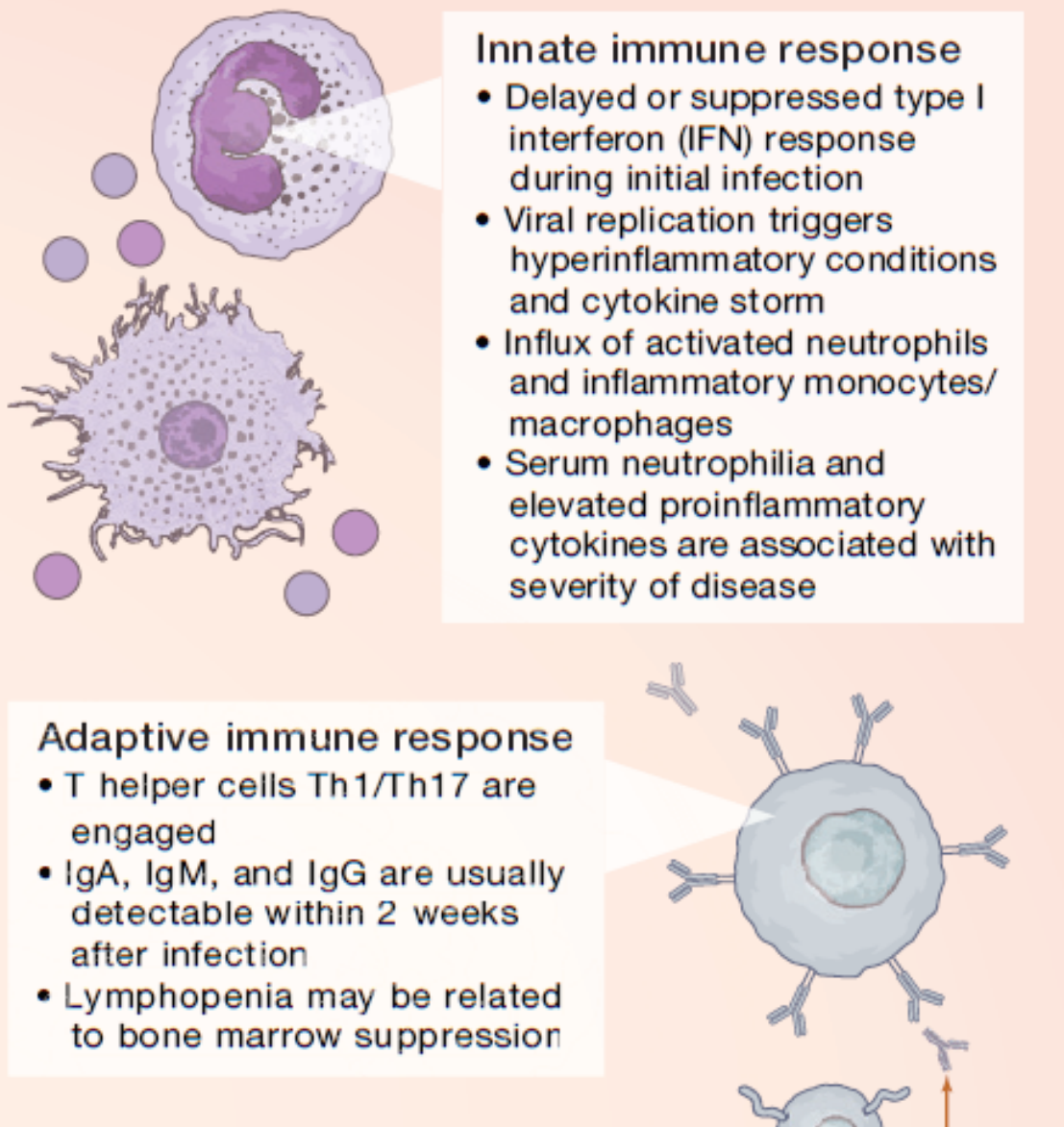
1. Ampliar la comprensión de la biología del virus

2. Permitir el seguimiento de los diagnósticos, vacunas y tratamientos.

- Seguir los cambios en la antigenicidad del virus en respuesta a la introducción de vacunas
- Seguir la aparición de mutaciones de resistencia y medicamentos.
- Identificar los marcadores genéticos del virus asociados con la gravedad clínica para priorizar las intervenciones.

3. Mejorar nuestra comprensión de la epidemiología y la transmisión del SARS-CoV-2.

Mutaciones en epítomos/antígenos



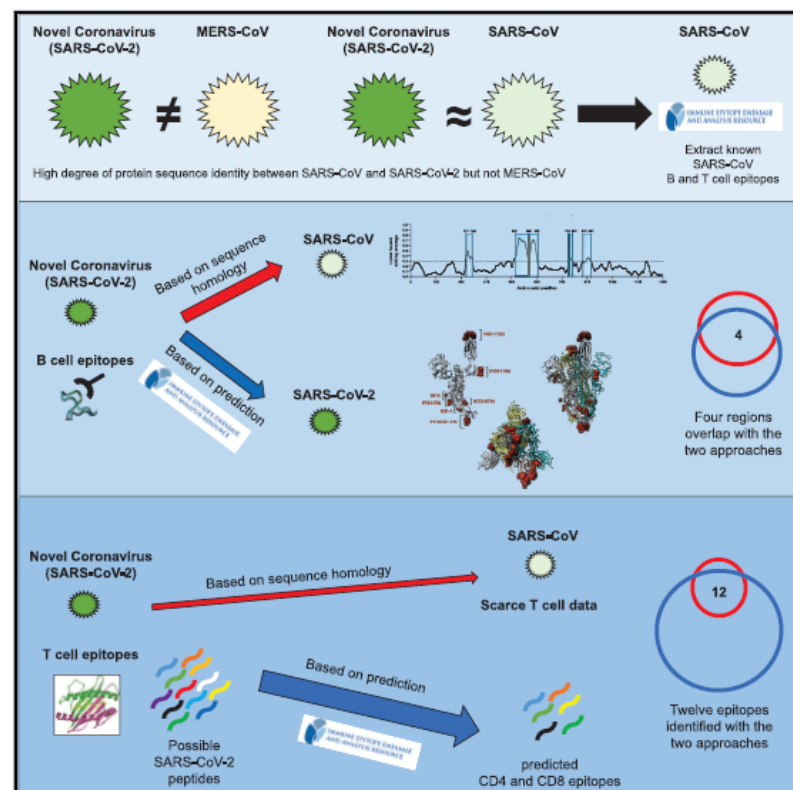
Cambios en péptidos del virus reconocidos por el sistema inmune **podrían cambiar la antigenicidad y la efectividad de una vacuna**

Mutaciones en epítomos/antígenos

Cell Host & Microbe

A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2

Graphical Abstract



Authors

Alba Grifoni, John Sidney, Yun Zhang, Richard H. Scheuermann, Bjoern Peters, Alessandro Sette

Correspondence

alex@lji.org

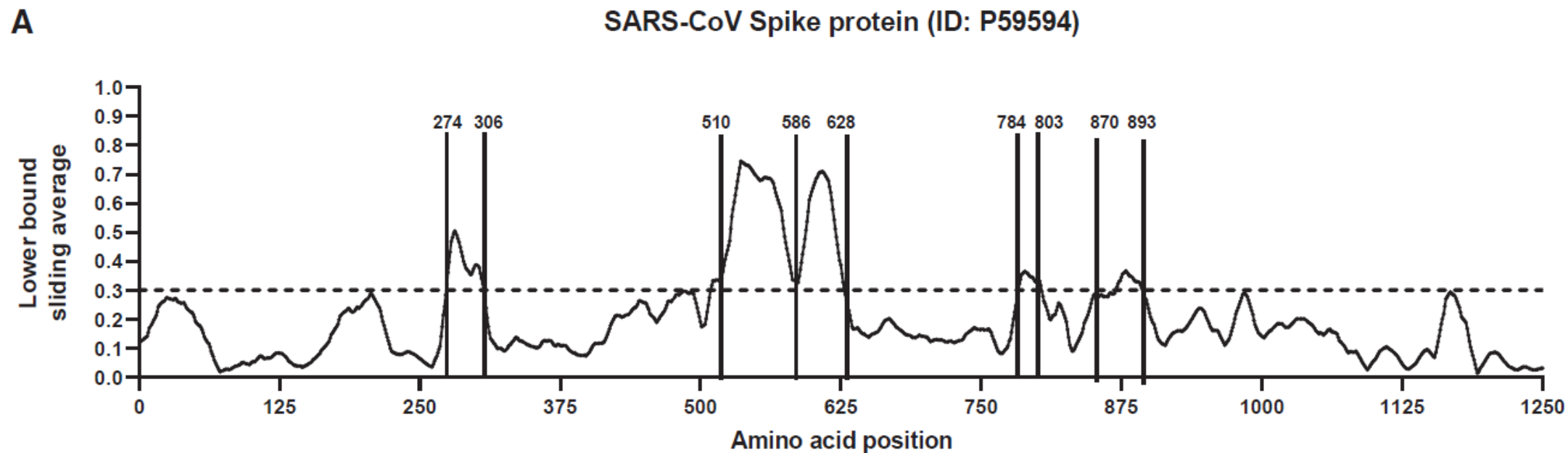
In Brief

Grifoni et al. identify potential targets for immune responses to the 2019 novel coronavirus (SARS-CoV-2) by sequence homology with closely related SARS-CoV and by *a priori* epitope prediction using bioinformatics approaches. This analysis provides essential information for understanding human immune responses to this virus and for evaluating diagnostic and vaccine candidates.

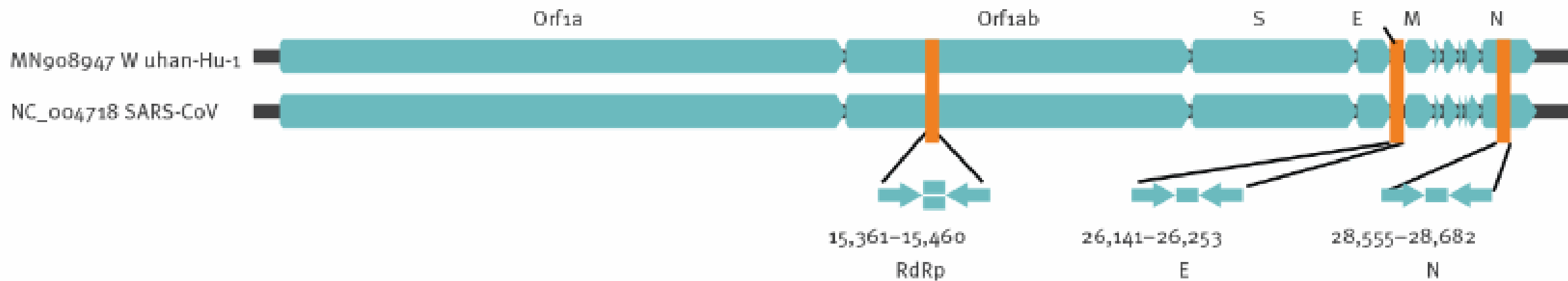
Predicción epítomos en SARS-CoV2

Mutaciones en epítomos/antígenos

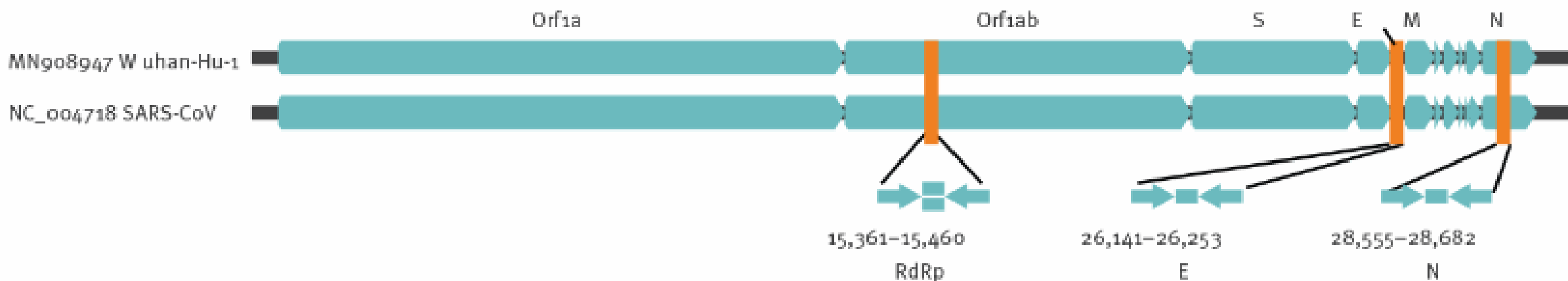
Regiones inmunodominantes



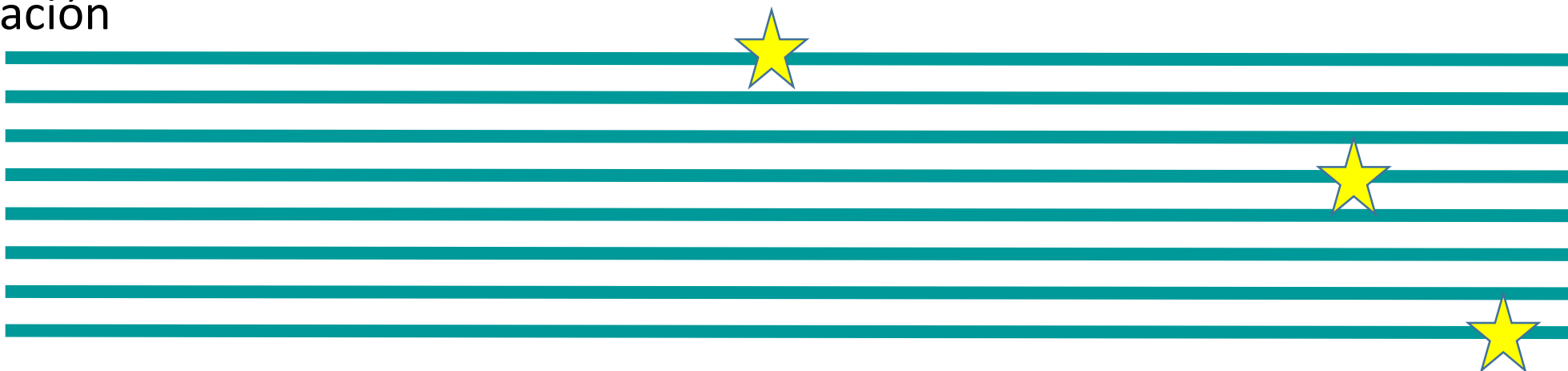
Impacto mutaciones en diagnósticos



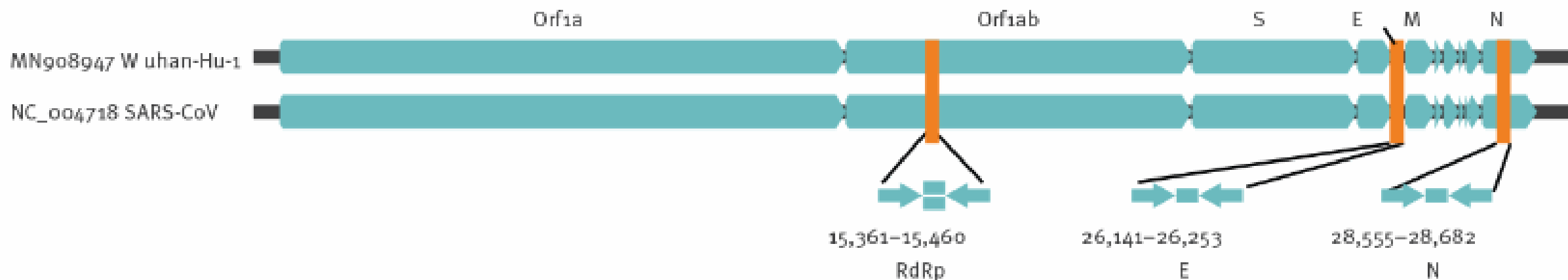
Impacto mutaciones en diagnósticos



Secuencias población



Impacto mutaciones en diagnósticos



Algunos cambios en los genes detectados por los métodos de diagnóstico molecular tipo PCR **podrían** reducir la capacidad para detectar el virus

Epidemiología genómica

1. Ampliar la comprensión de la biología del virus
2. Permitir el seguimiento de los diagnósticos, vacunas y tratamientos.
- 3. Mejorar nuestra comprensión de la epidemiología y la transmisión del SARS-CoV-2.**
 - Evaluar el efecto de las intervenciones no farmacológicas (como restricciones de viaje, sociales distanciamiento y aislamiento) en la propagación del SARS-CoV-2.
 - Inferir la contribución relativa de la transmisión local versus los casos importados.
 - Distinguir cadenas de transmisión individuales a nivel local.
 - Tasas estimadas de crecimiento epidémico: número reproductivo.
 - Determinar patrones de propagación de virus entre regiones y entre subgrupos de población.

Epidemiologia genómica

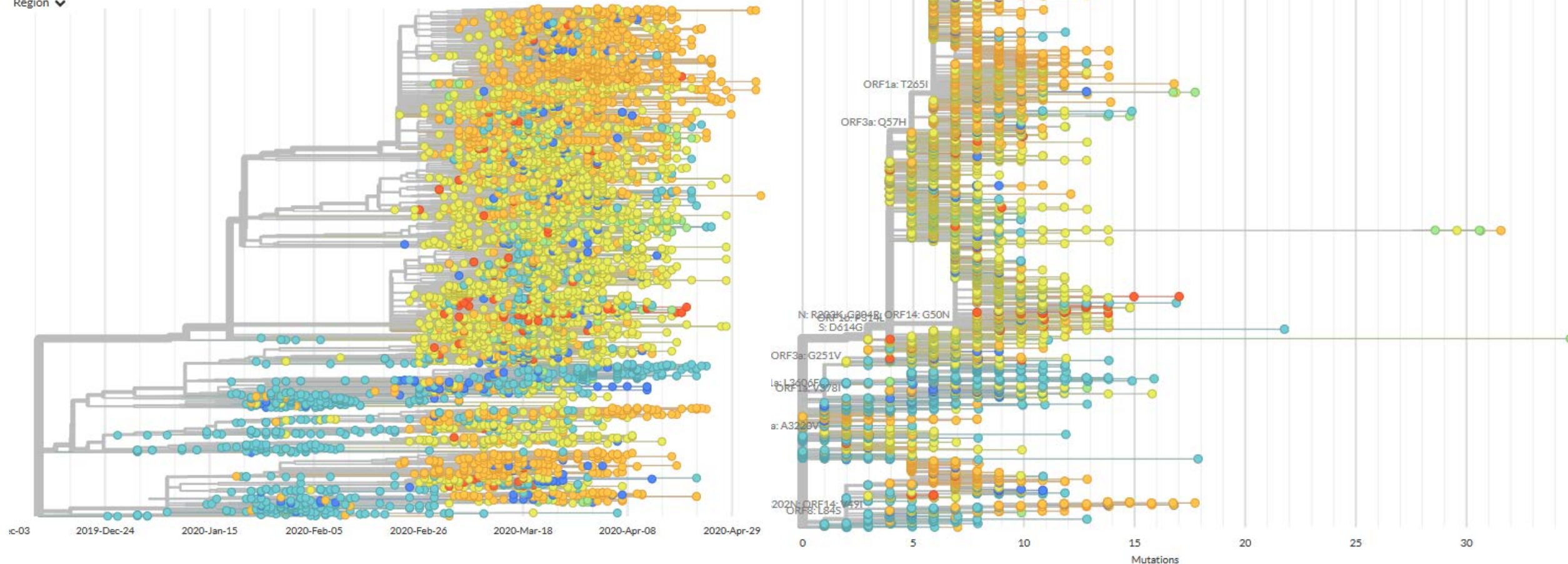


Nextstrain

Phylogeny

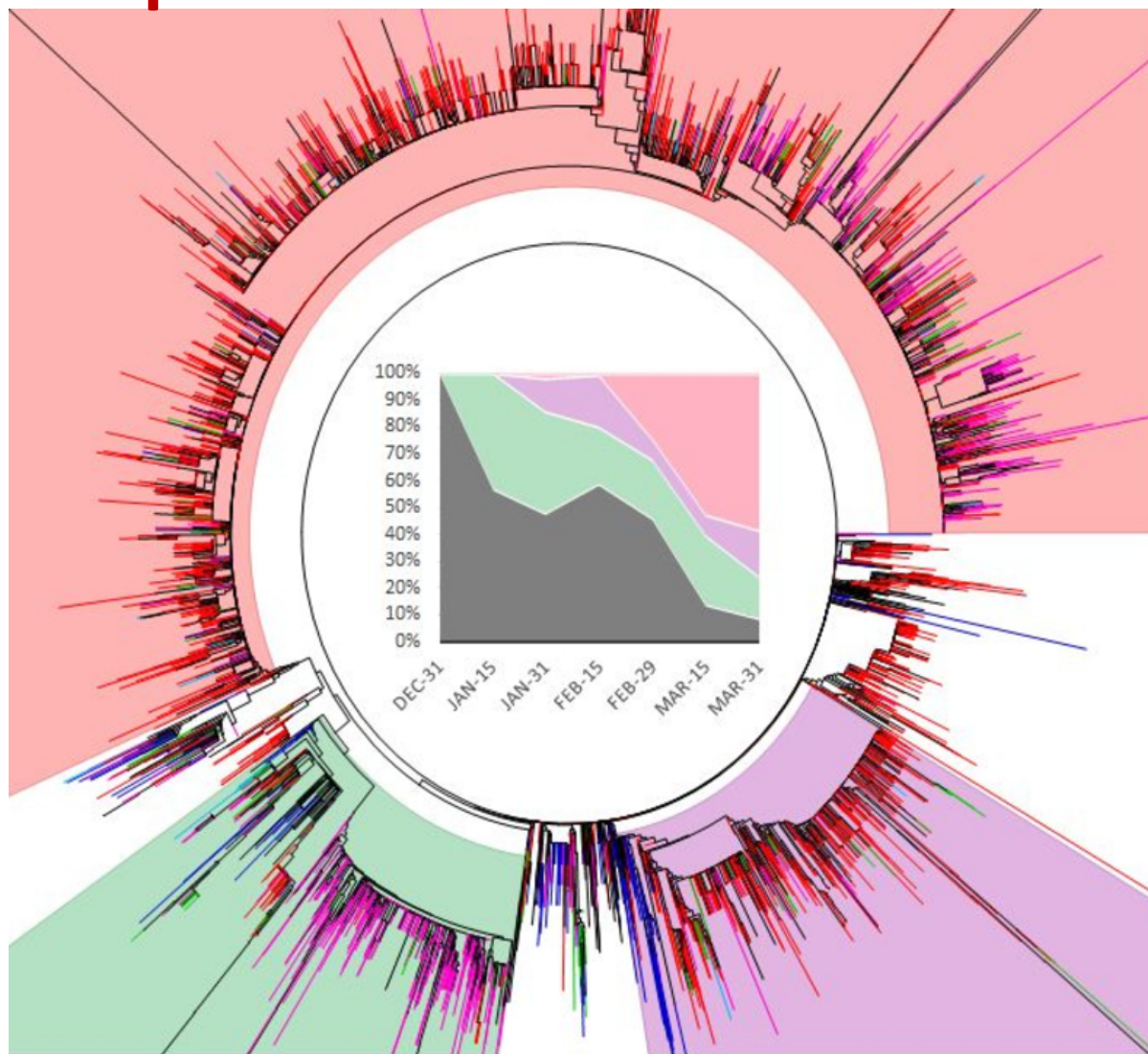
Region ▼

RESET LAYOUT Region ▼



Grupos de secuencias

Grupo G: S-D614G



Notable changes:

8,458 full genomes (+964)

(excluding low coverage, out of 9,076 entries)

S clade 1,026 (+48):

21 Spain, 15 Scotland, 5 USA/WI, 2 England, 2 USA/MI, 1 Germany, 1 USA/NC, 1 Korea

G clade 4,850 (+606):

227 Scotland, 214 England, 65 USA/NY, 30 USA/WI, 25 Spain, 19 Germany, 9 Belgium, 8 Taiwan, 7 USA/MI, 1 France, 1 USA/NC

V clade 1,202 (+159):

98 England, 57 Scotland, 2 USA/NY, 1 Korea, 1 Belgium

Other clades 1,380 (+151):

73 Scotland, 64 England, 10 Belgium, 2 Germany, 1 Sweden, 1 Czech Republic

Grupo S: ORF 8-L84S

Grupo V: NS5B-G251V

Epidemiologia genómica



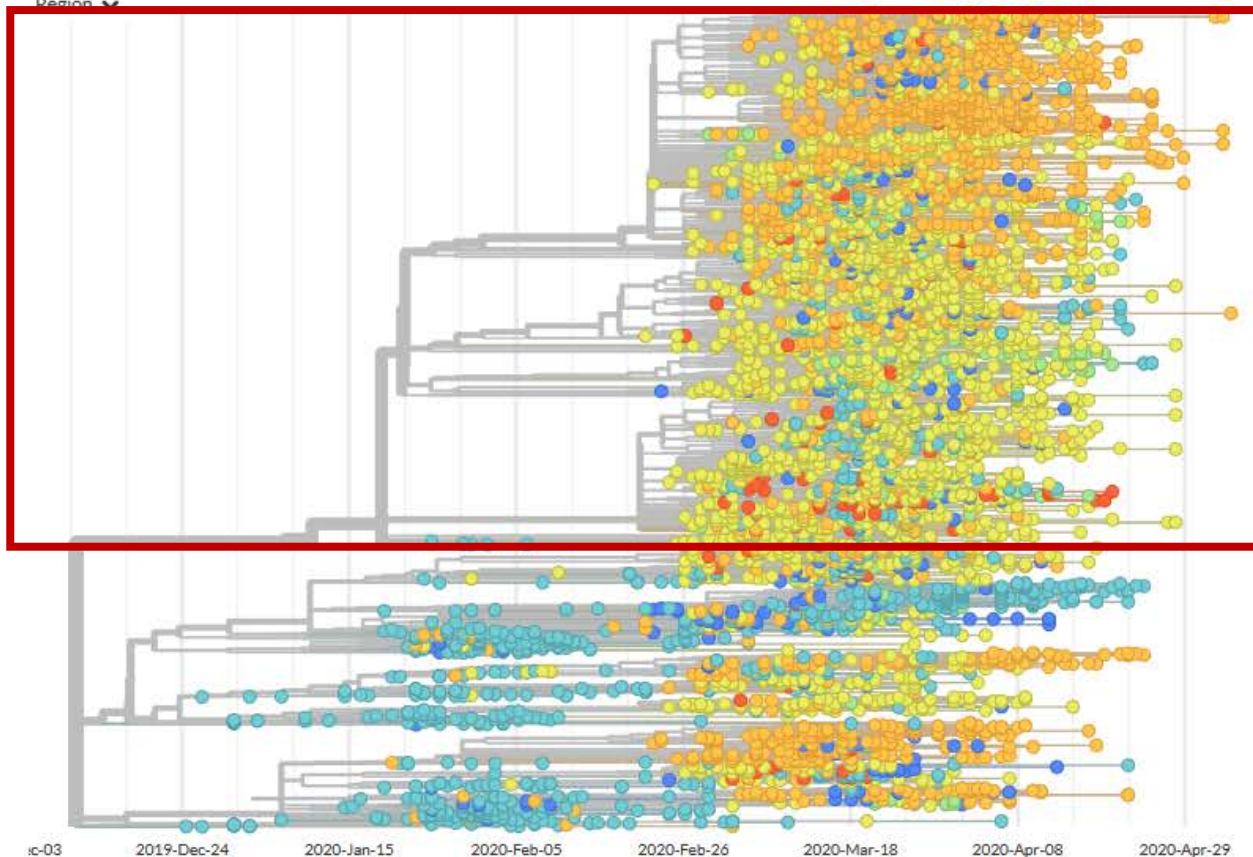
Nextstrain

GISAID

Phylogeny

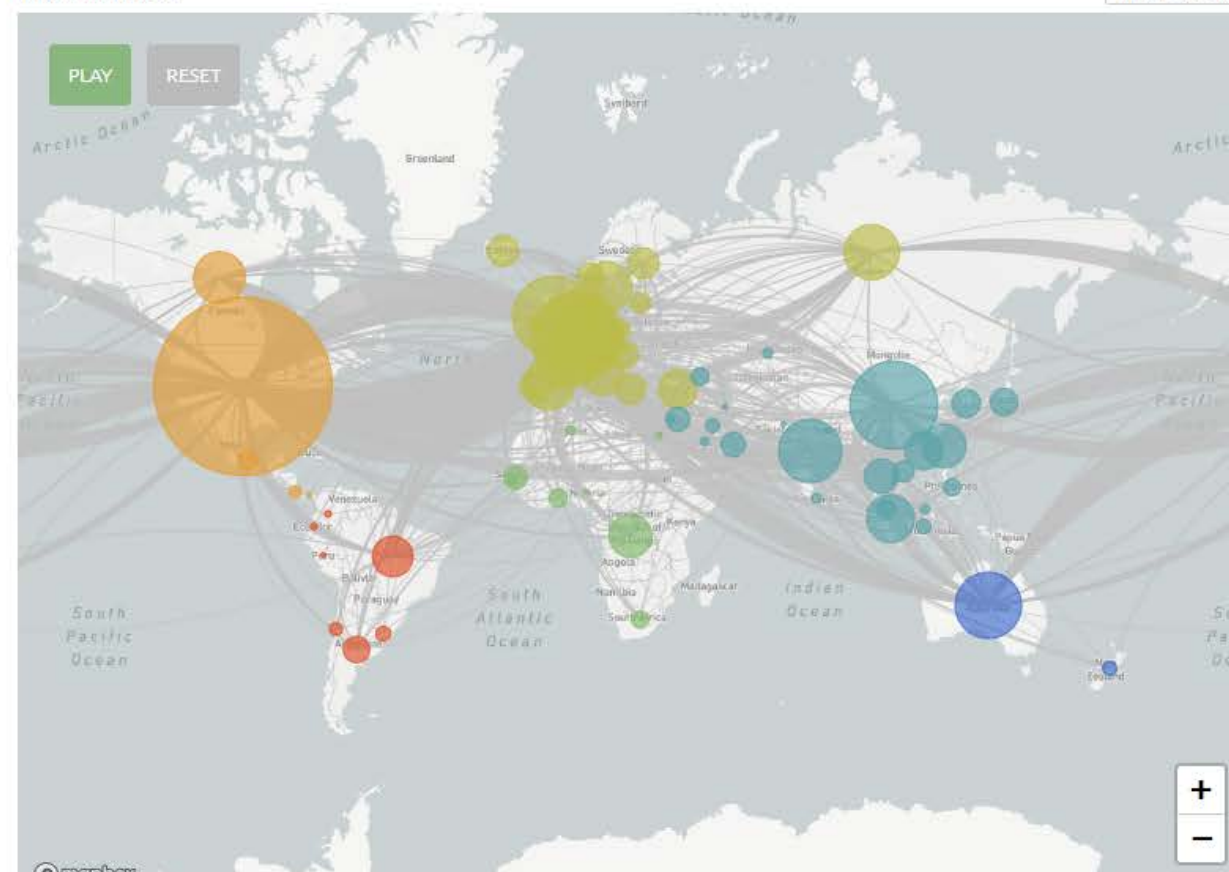
Region ▼

RESET LAYOUT



Transmissions

RESET ZOOM



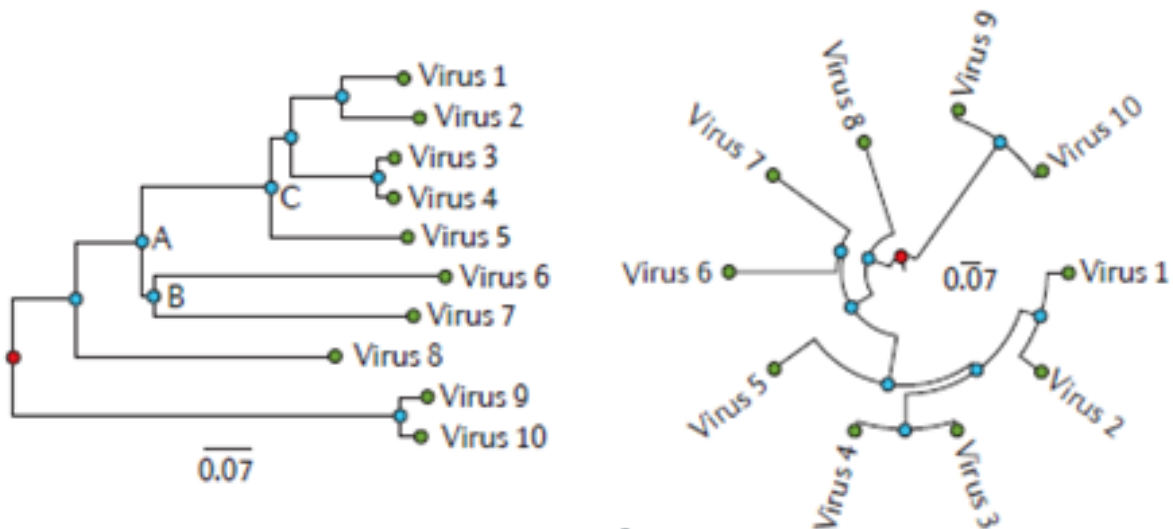
Epidemiología genómica

Cadenas de transmisión

AGACTAATTGCGCT
AGACTAATTGCGCT
AGACTAATTACGCT
AGACTAATTGCGCT
AGACTAATTGCGCT

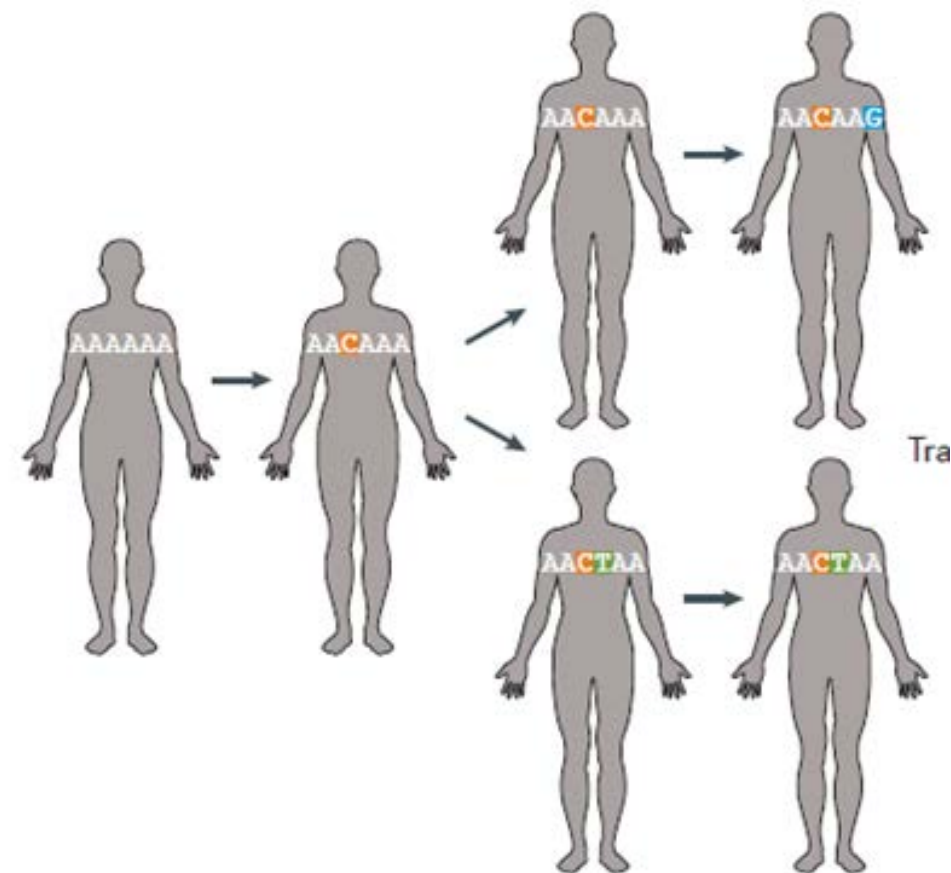
Identification of high-quality variants

Comparative analysis and visualization of relationships



Phylogeny inference

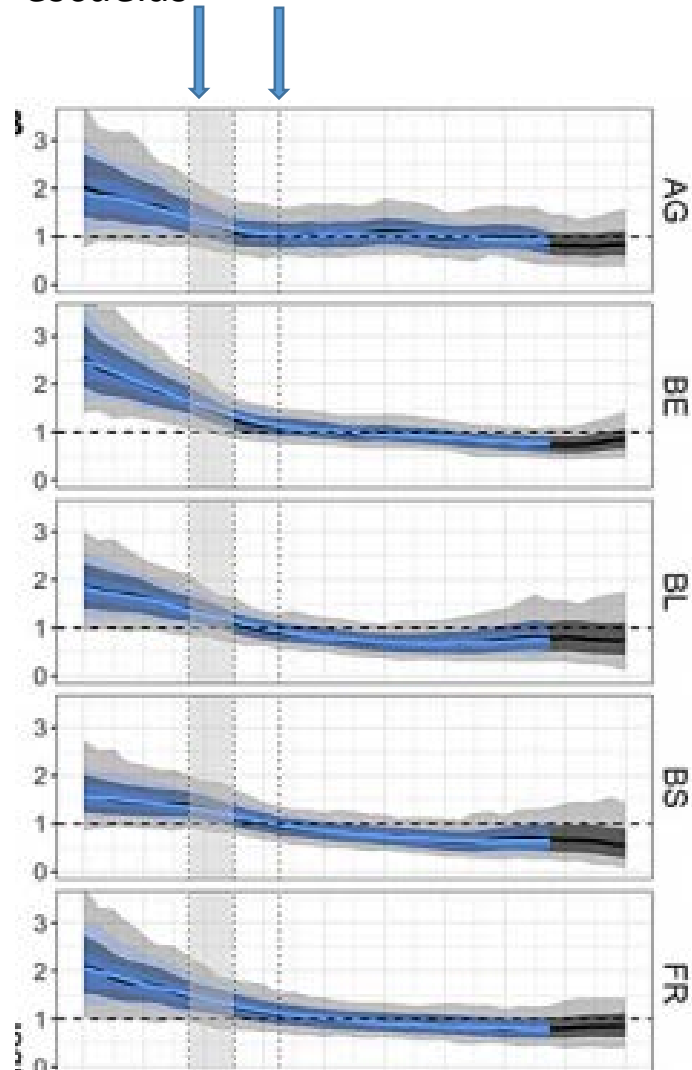
Inferencia de la transmisión



El número reproductivo a través del tiempo y su cambio con las intervenciones de salud pública

Cierre de tiendas y
escuelas

Prohibición de reuniones
de más de cinco



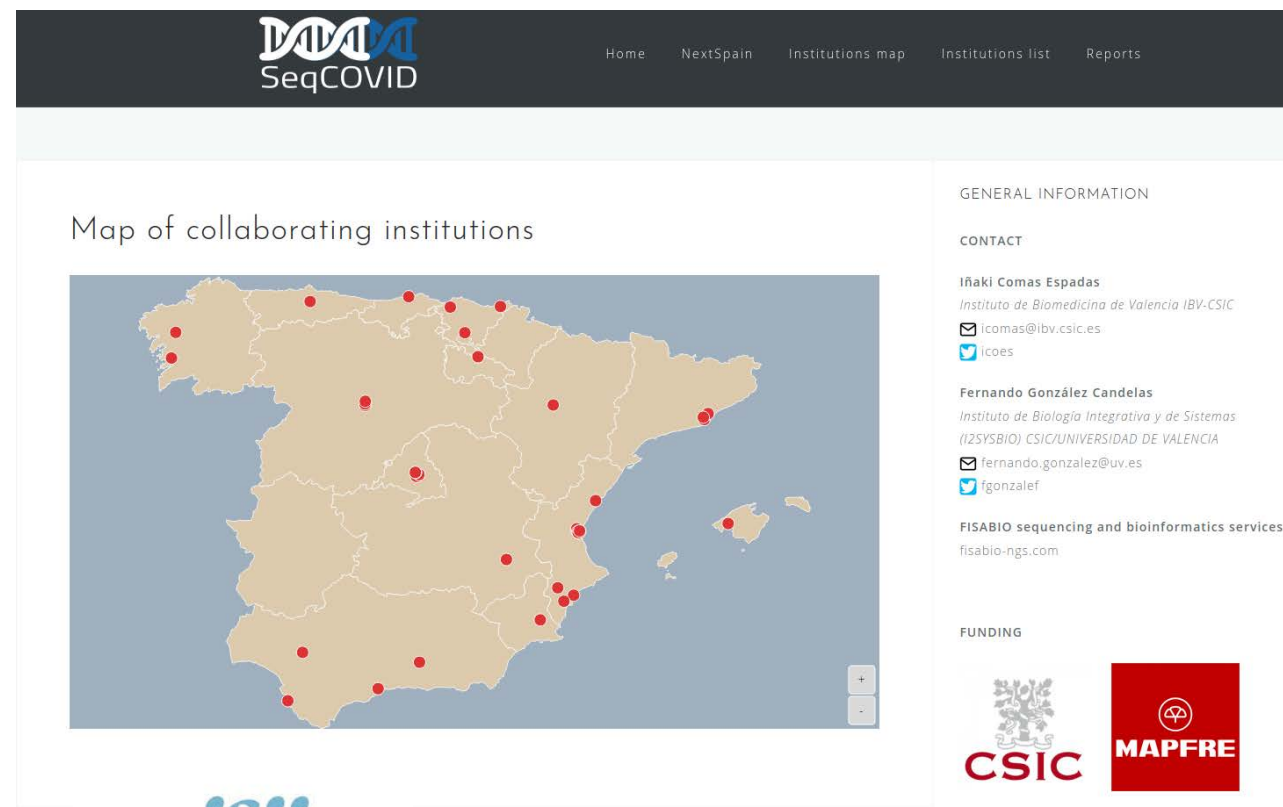
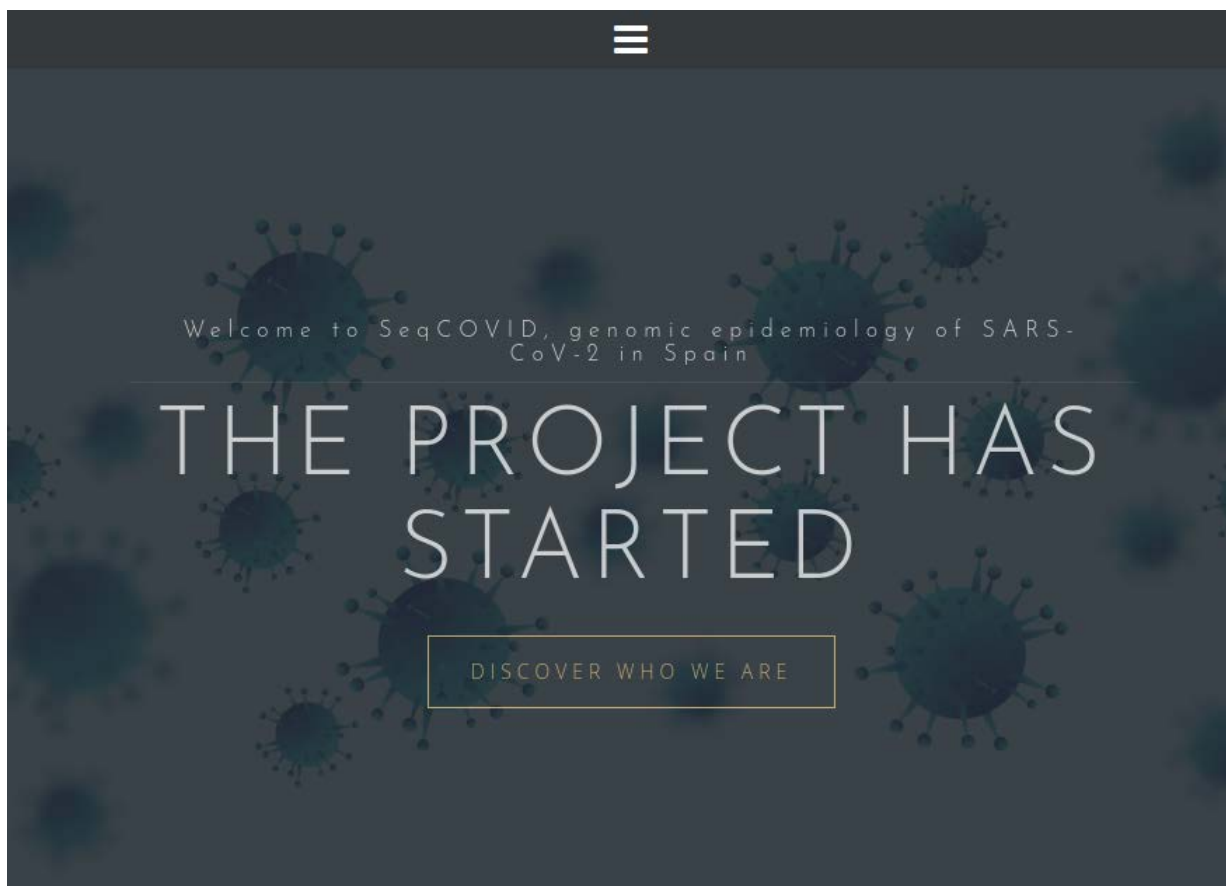
Descenso de R_t después de las intervenciones

Solo revela una correlación entre una disminución de $R(t)$ y una implementación de medidas, pero no una causalidad.

Sesgos en los análisis derivados de cambios en los esfuerzos de detección a lo largo del tiempo

SeqCOVID: Genomic epidemiology of SARS-CoV-2 in Spain

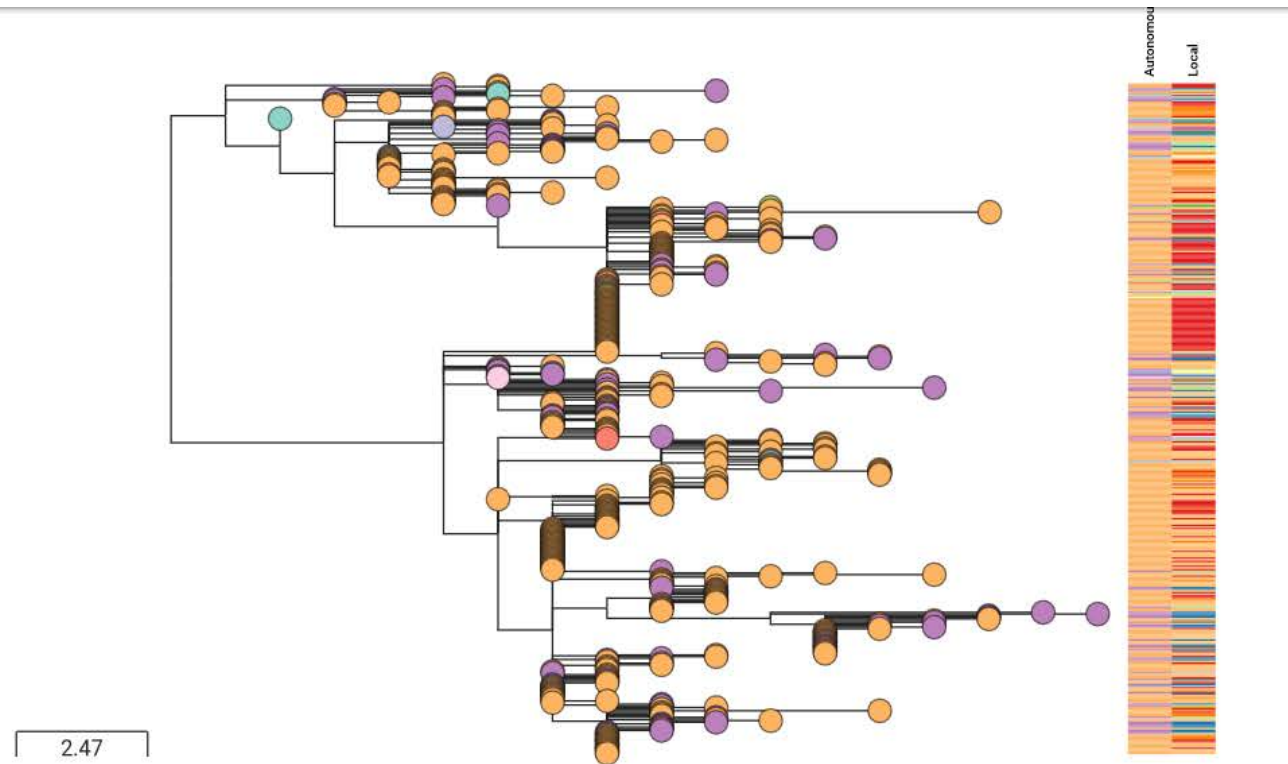
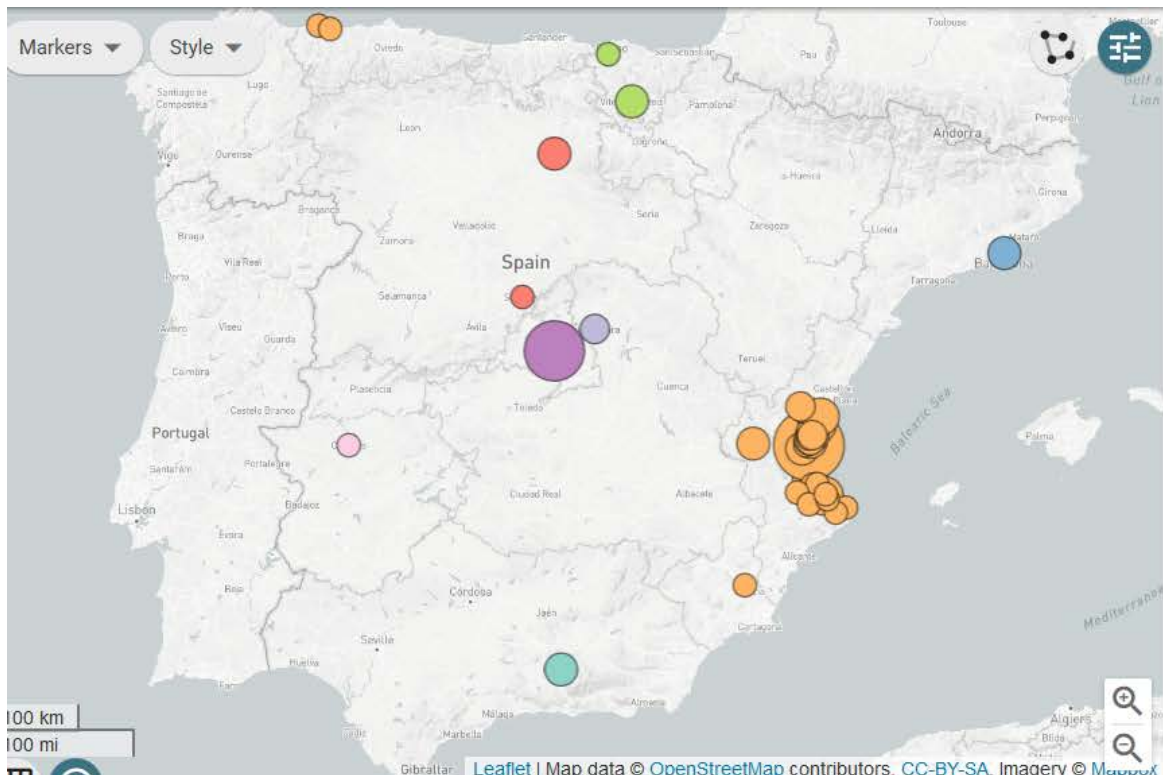
- Goal: 20,000 seqs along the 1st year
- Completed: 600 sequences
- More than 40 Spanish hospitals
- Based on diagnostic samples remnants
- Incorporating epidemiological information



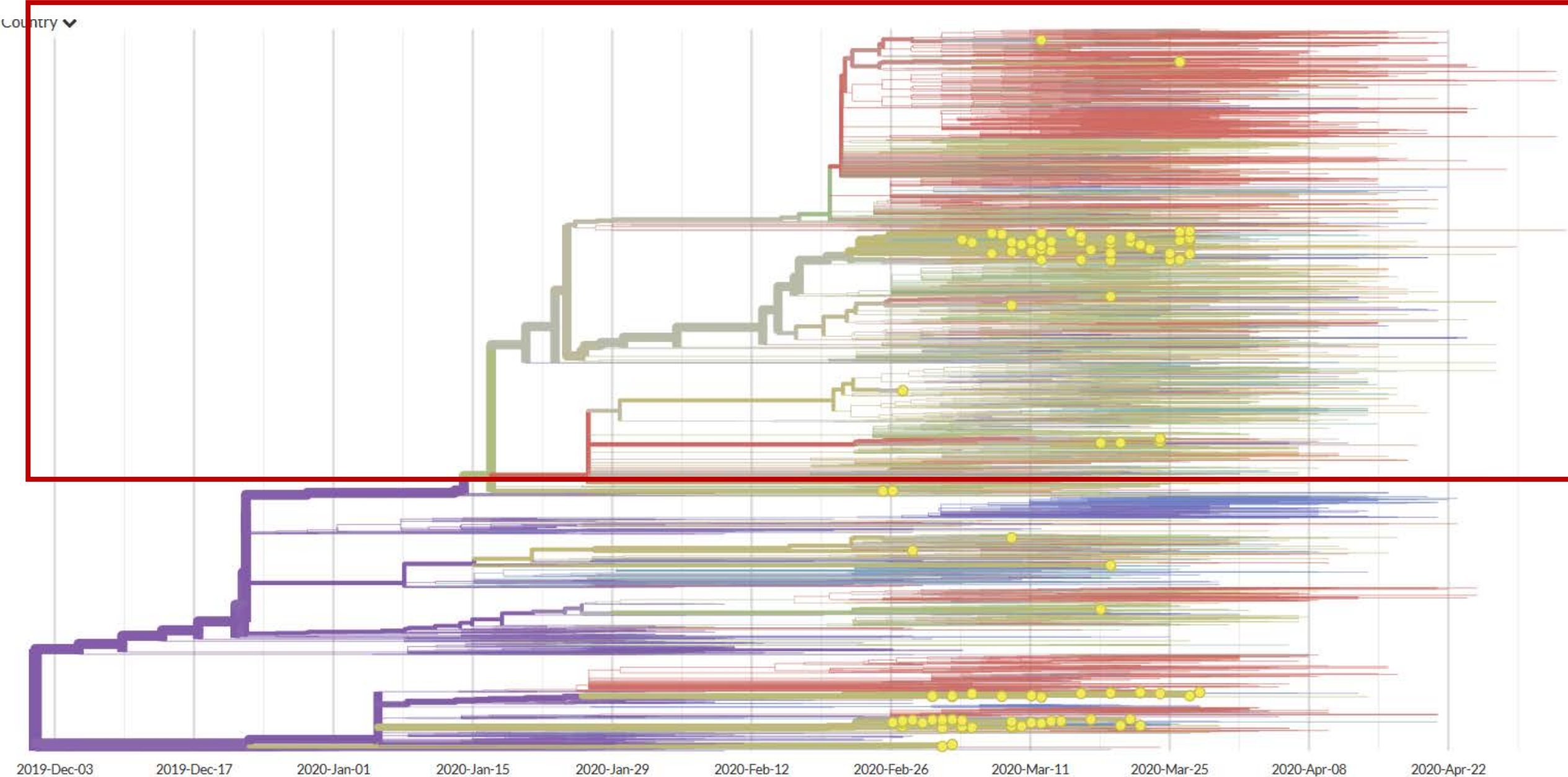
Secuencias analizadas

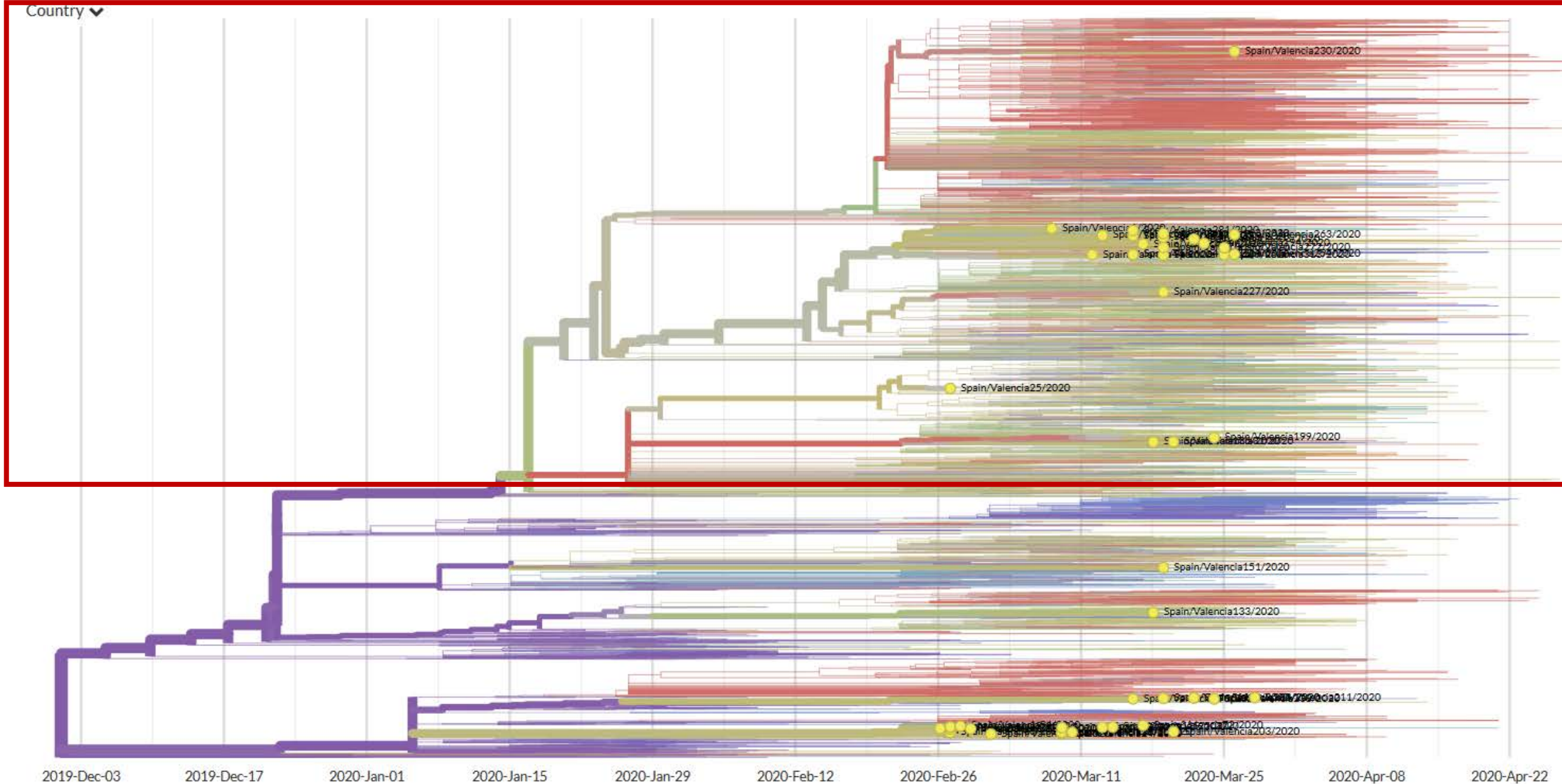


Introducciones múltiples en España



Introducciones múltiples en España





Circulación de SARS-CoV-2

Los brotes están relacionados dentro de cada país pero también entre diferentes países.

El virus se ha introducido y reintroducido varias veces, incluso en lugares donde la primera ola de infecciones ya pasado.

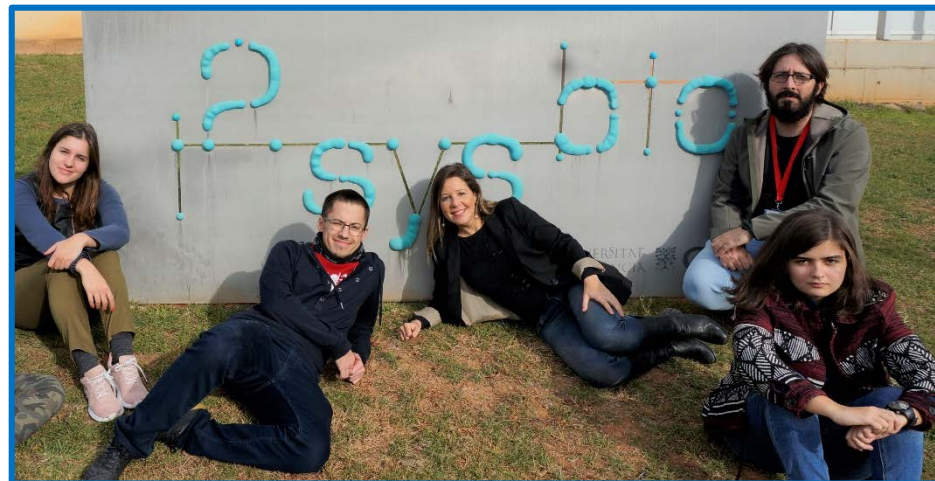
Genómica de
la tuberculosis



Secuenciación
FISABIO



PathoGenOmics



EpiMo

